

## Fusion on shifting foundations

**Creating affordable energy by fusing deuterium and tritium is still only a dream, while costs of fusion research continue to climb. The community's plans still command political support, but need bolstering by a fundamental review of activities.**

SINCE its launch in 1988, the collaboration developing the International Thermonuclear Experimental Reactor (ITER) has been the principal focus of political debates about fusion. Those debates have themselves been dominated by issues of management, costs and multinational politics. But the signs are that the foundations of the project are shifting: last month's drastic reduction in the US fusion budget, Russia's chronic shortage of resources, recent nervousness expressed by European research ministers — only Japan remains bullish.

Even Japan's enthusiasm (\$365 million is spent annually on fusion) carries risks. It is based partly on concern about the future of the country's energy supply, but nobody's vision of the technology can be complete enough for fusion yet to be responsibly offered as a likely solution. The most powerful motivators appear to be the vested interests of heavy-engineering industry and of bureaucrats within a system that, by comparison with the West, is short of mechanisms to apply the brakes to high-technology projects. But if the costs rise and budgets tighten, such backing may not be enough. Japan's history of technology development is exemplary, but only if it is successful in a bid to host ITER might its commitment become effectively guaranteed.

The US fusion energy research programme is being cut by one-third this year, entailing a painful adjustment that will probably include the closure of the largest US fusion machine, the Tokamak Fusion Test Reactor (TFTR) at Princeton, New Jersey. This retrenchment results from budgetary pressure aligned with scepticism about fusion in particular, but not from any fundamental re-evaluation that might justify such a significant cut in funds. Nevertheless, it does in effect represent a confrontation by Congress of issues that the US administration and its ITER partners have so far ducked. The cut in the US programme from \$365 million last year to \$244 million is the latest stage in a steady decline which had already seen the scope of the effort halved since late-1970s. Europe, meanwhile, has been a redoubtable supporter of fusion — but its governments have recently begun to wobble.

In public, at least, the US fusion community is rallying behind a plan presented two weeks ago to Martha Krebs, head of research at the Department of Energy, by the Fusion Energy Advisory Committee (FEAC). That plan calls for a slimmed-down programme whose emphasis is shifted away from technology and towards basic science. It would maintain the US commitment to the current engineering design stage of ITER (due to be completed in 1998) and offer more support for basic science proposed by a wider base of independent, university-based investigators. But there will be no money for the extra materials science that FEAC says is needed. That is unfortunate, because the lack of materials that can survive tokamak conditions for any length of time is a critical impediment to fusion's viability.

The cuts are placing immense strains on the US community. The hope must be that Congress will perceive fusion scientists as adequately "united" behind the FEAC plan and, thus encouraged, give them \$250 million next year to implement it. FEAC calls that a "niche" programme, but the niche is large, embracing most of the important science that will underpin ITER or any similar machine. If that happens, reports of the death of US fusion will

have been grossly exaggerated.

The ITER agreement commits the partners only to producing sufficiently detailed designs by 1998 to allow construction "either alone or together". Nobody is reneging on that, but the total cost of the project is estimated by the US government at \$10 billion or more. Europe and Japan reacted angrily when the United States suggested a scaling down of ITER last summer, but subsequent European concerns serve only to amplify the undermining by the United States of confidence in ITER's future.

The unpalatable truth is that, because of changed circumstances, ITER's collaborating countries are now all, to some degree, unreliable. Between them they are spending more than \$1 billion a year on fusion, of which ITER is only a part. They need to reassure themselves that, as a whole, the world's fusion community is pursuing its science with a rational use of resources. The OECD's Megascience Forum has been established to implement precisely such an investigation, but has so far been confined to relatively uncontroversial topics. Either there or elsewhere, taking international stock of fusion would benefit everyone. □

## Heed chemical warnings

**The UK government should consult independent researchers on chemical weapons.**

ONE welcome by-product of the ending of the Cold War has been the prospect of an international treaty banning the production and use of chemical weapons. The wording of the treaty was finally agreed, after 20 years of negotiation, at the beginning of 1993, and the number of countries ratifying the treaty is rapidly approaching the 65 needed to bring it into force. One remaining issue of concern is that, in their efforts to ensure the strict enforcement of the treaty, governments may be tempted to introduce regulations that unnecessarily harm research in fields such as brain science that use highly toxic agents.

Such concern surfaced last week in a debate in Britain's House of Lords on the implementation legislation that has been drawn up by the UK government (see page 479). There is no serious opposition to the legislation itself. But several participants rightly pointed out the dangers to research of over-rigid controls.

There is, of course, an element of special pleading involved. Scientists may object to the extra red-tape that will now inevitably be involved in working with some chemicals. The extra vigilance, whatever its inconvenience, is an inevitable part of the price that has to be paid for controlling the spread of what has been called "the poor man's nuclear weapon".

But the sensitivities of researchers must not be ignored, not least because their goodwill is essential for maintaining confidence in the effectiveness of control procedures. The British government is faced with the chance of making this easier by agreeing to set up an external advisory committee to monitor its regulatory efforts. Such a committee would provide a means by which selected members of the scientific community could raise issues of concern to their colleagues. That opportunity should be grasped.



# NIH agrees to temporary by-pass of AIDS office in allocation of grants

**Washington.** The US National Institutes of Health (NIH) has made a preliminary decision to remove control over \$1.4 billion in spending on AIDS research during the current fiscal year from its embattled Office of Aids Research (OAR), and to channel the money directly to its institutes instead.

A spokesperson for Harold Varmus, the director of NIH, said on Monday that the biomedical agency took this decision last week. "We believe it was the best reading of the law," says Anne Thomas, although adding that Varmus was still "behind the OAR and its authority".

The decision was taken by NIH officials after consulting with the Department of Health and Human Services (HHS), of which the NIH is a part. NIH officials say that the agency, four months into the financial year and facing thousands of researchers still waiting to receive grants after two government shutdowns, was forced to take the decision to get its funding back on track.

But Clinton administration officials are still insisting that the decision is not final, and that the OAR will ultimately keep its responsibility for distributing to the NIH's 24 institutes and centres what amounts to 12 per cent of the NIH's \$11.9 billion 1996 budget. "We are not yet done with this issue," said a spokesman for the White House's National Aids Policy Office, adding that President Clinton is "absolutely and firmly" committed to maintaining the budgetary responsibilities of the AIDS office.

But Anthony Fauci, head of the National Institute of Allergy and Infectious Diseases, the top recipient of AIDS research money

receiving about 42 per cent of the total, said that for the present the OAR will be by-passed. "The decision at this point is to go ahead with [circumventing the OAR]," he said last Friday, 2 February. The director of the OAR, William Paul, could not be reached for comment, and his office did not return telephone calls this week.

AIDS activists complain that the decision, if made final, would emasculate the 33-member office, responsible for overseeing and coordinating AIDS research at the NIH. Its most immediate impact, they say, would be to consign to obscurity the recommendations that are due to emerge shortly



**Varmus: AIDS research review expected soon.**

from a broad review of the NIH's AIDS research.

The study, which is to be released in March, has involved over 100 scientists and activists, and is expected to recommend fundamental changes in AIDS research at NIH. If OAR's authority to direct

money is removed, "the report becomes a toothless tiger you can bury in a library somewhere", says one researcher.

The current controversy over the role of the AIDS office stems from a vote by the Republican-dominated House of Representatives in August to remove OAR's role in directing research funds to the various NIH institutes — a role that individual institute

directors are known to resent.

A companion measure in the Senate retained the office's funding role. But neither bill became law, while a temporary measure approved recently, funding the NIH up to the end of September, is ambiguous as to whether the AIDS money should be distributed through the OAR, or go directly to the institutes.

While Varmus's office indicated that the NIH had decided on circumventing the OAR, officials in HHS and the Office of Management and Budget say that they are still working with the NIH to interpret the temporary spending measure, and to reach a final decision on whether to by-pass the OAR. "The administration is committed to an independent office of AIDS research, and is working to fix the perception problems that have arisen," says Victor Zonona, a spokesman for HHS.

AIDS activists say they were assured last week that the administration would not allow the OAR to be by-passed. Mark Harrington, a spokesman for Treatment Action Group, a New York AIDS activist group, says the group was "shocked" by the NIH's preliminary decision. He claimed Varmus had promised "many times" that support for OAR was "one of his highest priorities".

But Republicans in the House say the OAR should lose its funding role. A spokesman for John Porter (Republican, Illinois), chair of the appropriation subcommittee responsible for the NIH budget, says the full committee believes that "science, not politics, should make all judgements about scientific research". His subcommittee was largely responsible for boosting the NIH's 1996 funding by 5.7 per cent, which some feel has made Varmus loathe to risk alienating Porter by protecting the OAR.

The relevant bill already approved by the House specifies that AIDS research money must be funnelled directly to NIH institutes, the top two recipients being the National Institute of Allergies and Infectious Diseases and the National Cancer Institute.

Some point out that by-passing the OAR contravenes 1993 legislation that put the power to disburse AIDS research money in the hands of the office, which is strongly supported by AIDS activists. They say that it plays a crucial role coordinating AIDS research and avoiding duplication and waste. But conservative Republicans call the office, which was founded in 1988, a political interference in science by liberal Democrats who count the bulk of AIDS patients in their own constituencies.

**Meredith Wadman**

## Rockefeller biologist tipped for the top at Salk

**Washington.** James Darnell, a prominent molecular biologist who is head of the laboratory of molecular cell biology at Rockefeller University, New York, is being tipped to succeed Francis Crick as president of the Salk Institute in San Diego, California.

According to colleagues at Rockefeller, Darnell has been offered the Salk presidency, and is expected to decide whether to accept it this week. But neither the Salk Institute, nor Darnell himself, would confirm that the offer had been made. Anita Weld, a spokeswoman for the Salk, would say only that a search for a president is still in progress, led by Frederick Rentschler, vice chairman of the Salk board of trustees.

Crick, who won the Nobel prize with James Watson for the discovery of the double helical structure of DNA, stood down

last August after undergoing a heart bypass operation, but remains on the faculty there.

Sources at Rockefeller say that Darnell, who is best known for his work on cell signalling and gene expression, was meeting this week with Torsten Wiesel, president of the university, before deciding whether to make the move. Many hope he will turn down the offer; Darnell has been responsible for managing junior faculty at Rockefeller, and 20 of them have written to Wiesel asking him "to do everything possible to keep Jim here at the university".

The Salk Institute, which was established by Jonas Salk, the polio vaccine pioneer, in 1960, now has 50 faculty members and a total of 500 staff engaged in biomedical research, especially in molecular biology, genetics and neuroscience. **Colin Macilwain**

# European space science braces for mission cuts

**Paris.** The European Space Agency (ESA) is expected to announce major cuts in its space science programme later this month, including a 10 per cent reduction in the budgets of all its science missions, the cancellation of at least one planned mission and delays to others.

The cuts are part of a drive to reduce costs following last year's decision by ESA's 14 member states to freeze spending on space science over the next three to five years. Also in doubt are plans to collaborate in upgrading the Hubble Space Telescope (see *Nature* **368**, 786; 1994) and to continue operating satellites beyond their scheduled design lifetimes.

This was the gloomy news that emerged from a meeting last month of ESA's Space Science Advisory Committee (SSAC), whose conclusions are expected to be endorsed by its Science Policy Committee at a prognosis meeting next week.

ESA's Horizon 2000 science programme began in 1985 and is scheduled to run until 2005. Four large 'cornerstone' missions, each costing around ECU625 million (US\$795 million) make up the backbone of the programme, which also includes several 'medium-sized' missions at around ECU345 million each. The quality and cost-effectiveness of the programme has been widely acclaimed, and the same structure will be used within Horizon 2000-plus, a series of missions scheduled from 2006 onwards.

The proposed cuts in Horizon 2000 result

from the decision taken at last year's ESA space ministers' meeting in Toulouse to freeze annual spending on space science at ECU347 million annually from 1996 to 1998, when a decision will be taken on whether to extend the freeze until 2000 (see *Nature* **377**, 667; 1995).

ESA had previously proposed that the budget for Horizon 2000 should be increased at the rate of inflation over the next five years. But this was opposed by both Germany, which argued that the budget should be kept constant, and the United Kingdom, which argued that the agency would be able to cut the costs of its science missions by 25 per cent over this period without reducing their scientific value.

But Roger Bonnet, head of science at ESA, claims the SSAC proposals confirm his prediction that cuts of such magnitude could not be made without jeopardizing the science programme. The proposed 10 per cent cut in all science missions will inevitably increase the risk in the missions, says Bonnet, as they could only be achieved by eliminating parts or using cheaper alternatives. The next medium-sized mission, which is scheduled to be selected in June (see *Nature* **369**, 90; 1994), will be used as a pilot test to decide the possibilities for savings, he says.



**Friend or foe? Space station is putting pressure on budgets, causing dissension in countries involved (right and below).**

Bonnet also argues that the potential for savings is narrowed by the need to budget for unpredictable events. He points out that the delay to the Ulysses mission, caused by the explosion of the space shuttle Challenger in 1986, cost ESA ECU100 million. Similarly, the costs of the infrared space observatory satellite, which was launched last year, overran because of a two-year delay in the mission caused by technical problems, and an under-estimation of running costs.

But Bonnet's complaints have so far cut little ice with those who continue to believe that ESA can reduce its operating costs considerably without significantly affecting its science programmes. Ian Corbett, for example, director of science at the UK Particle Physics and Astronomy Research Council (PPARC), remains convinced that ESA can restructure its programmes "without any loss of science".

## Heads roll at French space agency

**Paris.** The French government last week announced that it had accepted the resignation of André Lebeau, the president of the French space agency, CNES, and has terminated the contract of Jean-Daniel Lévi, the director-general of the agency. Lebeau's successor is Alain Bensoussan, the current president of the national computer technology research agency INRIA; Lévi has not yet been replaced.

Two factors are said to lie behind the reshuffle at the top of CNES. One is the government's reported impatience with a power battle between Lévi, who was appointed in 1990, and Lebeau, who arrived in January 1995, having previously been president of Météo-France. Lévi was also an adviser to the late president François Mitterrand, and his position had been precarious since the election of Jacques Chirac as president.

But some observers feel that an addi-

tional reason for the reshuffle is that both Lévi and Lebeau have been lukewarm in their support for the international space station. Their differences with the government are said to have come to a head after it agreed to a large increase in its contribution to the station, at a meeting of Europe's space ministers in Toulouse last October (see *Nature* **377**, 667; 1995).

This decision, reportedly taken under pressure from the German government — Europe's largest supporter of the space station — put paid to Lévi's efforts to stop the escalation of CNES's contributions to the European Space Agency in order to reinforce its national and bilateral programmes. Moreover, in order to help pay for France's space station bill, CNES will have to make economies of 5 per cent a year in its planned budget over the period 1996–2000 (see *Nature* **378**, 5; 1995). **Declan Butler**

## US deal buys into

Washington. US and Russian space officials steered the international space station project through its latest political storm last week, agreeing on a plan that lets Russia meet its financial commitments to the project while keeping its own Mir station alive for several more years.

The United States will launch two more space shuttle flights to Mir, in addition to the seven that had already been planned. The additional flights, both in 1998, will carry supplies to the Russian station, which was to have been shut down next year but which will now be extended at least until 1999.

By picking up some of the cost of re-supplying Mir, the United States makes it more likely that its cash-strapped partner will be able to pay for the launch vehicles needed to help build the space station. In exchange, the National Aeronautics and Space Administration (NASA) will get more flight time on board Mir for its own astronauts and experiments.



Corbett says that Bonnet is "facing the challenge positively", pointing out that the proposal to carry out a pilot test is evidence that there is still room for improvement. But he contests Bonnet's claim that the proposed cuts are the direct consequence of the decisions taken at Toulouse, arguing that ESA's proposed mission schedule was already over-ambitious, and that the agency would have had to cancel or delay missions even without the ministerial decision.

Yet while PPARC officials claim that the SSAC proposals are simply a "recognition of this reality", Bonnet says that it is because Horizon 2000 is coming under increasing pressure to deliver more goods than originally planned that further cuts cannot be made without jeopardizing the programme's scientific content.

He points out that the programme is at the peak of its activities, and that ESA's budget is increasingly stretched by, for example, the need to meet the unanticipated demand of space scientists for the handling and storage of mission data. Similarly, researchers have been keen to exploit the possibilities of continuing to operate satellites, such as Ulysses and Giotto, that have survived beyond their scheduled design lives. But the cuts mean that ESA will probably no longer be able to fund proposals to prolong missions.

Bonnet's protests may seem like crying over spilt milk, given that the decision on the immediate cuts has already been taken. But his vigorous defence of ESA's science programme is based on the prospect of further cuts looming on the horizon. In particular, it is widely expected that Britain — and perhaps other member states — will push for an extension of the budget freeze from 1998 to at least 2000.

**Declan Butler**

## Italy 'cannot afford' space station role

**Munich.** The decision last year of Giorgio Salvini, the Italian research minister, to commit Italy heavily to the European Space Agency's participation in the International Space Station, has been strongly criticized by a panel of prominent scientists. The panel was set up last summer by the Italian government to report on Italy's national and international space programmes.

At the agency's ministerial meeting in Toulouse, France, last October, Salvini ensured Europe's continued involvement in the space station by confirming that Italy would provide 19 per cent of its costs (US\$318 million) until 2000. But in its report, published in Rome this week, the review panel says that Italy can neither afford such a contribution nor hope to receive significant scientific or industrial returns from the station, which it describes as being empty of any value apart from political prestige.

Members of the group that prepared the report include Antonio Ruberti, Italy's former research minister, Riccardo Giacconi, director of the European Southern Observatory, and Carlo Rubbia, the former head of the European Laboratory for Particle Physics (CERN), who originally advised Salvini not to commit Italy heavily to the programme.

In its report, the group acknowledges that the Toulouse decisions "cannot be ignored". But it proposes that Italy's contributions to the space station should be "reduced to the minimum possible level both during the development phase and

on the utilization phase, which is still very poorly defined". It also recommends that Italy "drastically reduces" its agreement with the United States to develop a pressurized logistic module for the space station, which could cost Italy IL535 billion (US\$339 million).

The report was commissioned to analyse the problems that had led to the current financial and administrative chaos in the Italian Space Agency ASI, to recommend ways of rationalizing the agency, and to propose an appropriate space policy for Italy. Its main argument is that a much too high proportion of the agency's budget of around IL1,000 billion a year goes to international space programmes, leaving national programmes underfinanced.

To correct the balance, says the advisory group, Italy's contribution to ESA should be reduced from around IL800 billion this year to around IL430 billion by 2000 by reducing its commitments to optional programmes, such as the space station and Earth observation and telecommunications programmes.

Speaking at a press conference on Monday, 5 February, Salvini hedged the issue of disagreement about participation in the international space station, saying that the agreements he had made at Toulouse were the "inevitable consequences of decisions taken at government level". But he acknowledged there was a need "to reconcile the decisions taken at Toulouse with those proposed in the report".

**Alison Abbott**

## Mir to keep Russia on board international station project

NASA will pick up costs that include \$124 million for one of the new shuttle flights — the second is being re-programmed from another mission already scheduled — as well as additional costs associated with flying more experiments. Exactly who will pay for what will be worked out at a further meeting in March between Russian and US space officials.

The agreement ends weeks of speculation that the partnership was in danger of falling apart. A proposal from Russia last autumn that the international space station should be built around existing elements of Mir, rather than from scratch, set off alarms both at NASA and in Congress, neither of which want another change of course when the station is due to start construction in November next year.

Daniel Goldin, administrator of NASA, vowed that the station design would not change, and two congressmen in charge of NASA funding, James Sensenbrenner (Republican, Wisconsin) and Jerry Lewis

(Republican, California), flew to Moscow last month to emphasize the point.

By that time, however, space station managers at the Johnson Space Center in Houston had already rejected the idea. They also vetoed a Russian proposal, presented at a meeting in Houston in mid-January, that the United States should pay Russia \$200 million over two years to help meet its obligations. "That [proposal] didn't stay on the table for very long," Wilbur Trafton, head of the space station project, said last week.

In contrast, sending additional shuttle flights to Mir "helps them, and it helps us", says Trafton. The two sides agreed to the plan at a meeting in Washington between US Vice President Al Gore and Russian Prime Minister, Viktor Chernomyrdin.

Although Russian participation in the station seems on a solid footing for the time being, NASA officials admit that future money problems could still jeopardize the partnership. As a hedge, Trafton says

engineers at the Marshall Space Flight Center in Alabama are working on contingency plans for building the station if Russia had to pull out.

But another agreement produced by the Gore-Chernomyrdin Commission may make that less likely. The number of commercial Western spacecraft allowed to use Russian launch vehicles is being increased from eight to fifteen. Revenue from those launches, and from experiments flown on Mir for paying customers, should improve Russia's financial position.

NASA, however, will now be looking for ways to pay for one more shuttle mission, as well as additional experiments on Mir, at a time when its budget is already overstretched. The strain is already beginning to show. Last week, Bryan O'Connor, director of the shuttle programme and a former astronaut, left the agency, apparently after disagreeing with a plan to cut NASA's costs by turning the shuttle over to a private operator.

**Tony Reichhardt**

# Japan reviews cost of nuclear reprocessing

**Tokyo.** Massive cost overruns in the construction of Japan's nuclear waste reprocessing facility in Aomori Prefecture in northern Japan have necessitated redesign and scaling back of the plant, pushing back completion by three years to 2003.

A joint announcement last month by Japan Nuclear Fuel Ltd (JNFL), the company that is building the facility, and the Federation of Electric Power Companies, revealed that the estimated cost of construction has doubled to ¥1,660 billion (\$US16 billion), further shaking public confidence in Japan's troubled plutonium policy.

Construction of the plant, which began at Rokkasho village in 1989, was originally expected to cost ¥840 billion, an estimate based on the construction cost of France's reprocessing plant at La Hague. JNFL claims that the increased cost is due to rising material prices, combined with the expense

of making the facility safe against earthquakes.

Despite doubling in price, the redesigned plant will be significantly smaller than the original. It will incorporate a single-stage process for separating uranium from plutonium rather than the original two-stage process, six rather than nine containers for high-level waste, and combined rather than separate facilities for evaporating and drying low-level liquid waste.

JNFL claims that the new design will have the same processing capacity as the original, namely 4.8 tonnes of plutonium per year from 800 tonnes of high-level waste. But industry observers such as Mika Obayashi of the Citizen's Nuclear Information Center in Tokyo are sceptical, given the reduction in the scale of the plant.

The cost overruns will be passed on to the power generation industry, and are expected

to cost the industry roughly ¥80 billion a year in increased reprocessing charges. Altogether, the cost of reprocessing fuel is calculated to add ¥1 to the cost of generating each kilowatt hour of electricity, which is already about 40 per cent more expensive than in France.

The effect of the cost overruns and delays is likely to add ¥2,000 billion (about US\$20 billion) to the total cost of establishing the nuclear fuel cycle within Japan, a central plank of Japan's long-term nuclear energy strategy. Japan currently has to use overseas facilities in France and the United Kingdom to reprocess its nuclear fuel.

This latest announcement adds to the woes of Japan's nuclear industry, already reeling after the shut-down in December of the experimental fast breeder reactor Monju in Fukui Prefecture in western Japan, and an attempted cover-up of the seriousness of the accident (see *Nature* 379, 196; 1996). This incident, which has brought into question both the technical feasibility of the FBR programme and the competence and trustworthiness of the administering bodies, has cast a shadow over the government's entire plutonium policy.

On the same day as the announcement on the Rokkasho reprocessing plant, the governors of three prefectures — Fukui, Niigata and Fukushima — visited the Prime Minister's office and the Ministry of International Trade and Industry (MITI) to urge the government to review the entire long-term nuclear energy programme. They expressed particular concern over the plan to burn mixed uranium-plutonium fuel (MOX) in conventional light water reactors (LWRs) from the late 1990s.

The three indicated that they will not permit the relicensing of reactors in their prefectures to burn MOX until the investigation into the Monju accident is completed. Almost two-thirds of Japan's 48 operating LWRs are in these three prefectures, including the two for which relicensing applications are expected to be made first. Neither the government nor the power utilities admits to any changes in the MOX policy.

Apart from domestic political opposition and popular distrust of the operating authorities, there is also concern among the international community about Japan's growing stockpile of plutonium. Despite the government's policy of not stockpiling plutonium — which could possibly be used to build nuclear weapons — its own figures show that plutonium stocks will rise from the present level of a few tonnes to about 70 tonnes by 2010, about half this figure representing imports of plutonium from Europe.

This estimate will no doubt be reduced because of the delays in the opening of the plant in Aomori.

**Stephen Barker**

## French genome pioneer goes private

Paris. Daniel Cohen, the French genome scientist, was due to announce this week that he is to resign as director of the Centre d'Etudes du Polymorphisme Humain (CEPH) in Paris, and merge his research group with the French biotechnology company Genset to create a world-class genome company.

Cohen, who will become general manager of genome research, says that Genset, with about 170 staff, will be the only European company large enough to compete with US rivals such as Human Genome Sciences, which has about the same number of staff.

Genset, which plans to seek a public flotation in the United States later this year, was created in 1989, and has since become a leader in the manufacture of oligonucleotides. It also carries out research in gene regulation, gene sequencing and computer analysis. Cohen's arrival will allow it to develop a comprehensive approach to genomics "from research to drugs", says Pascal Brandys, its chief executive officer.

In particular, Genset's business strategy will be to identify genes and use them to develop small-molecule therapeutics for the treatment of a broad range of diseases including cancer. The development of therapeutics would be carried out in collaboration with large drug companies.

The attraction of Genset, says Cohen, is its "industrial culture" and high-technology approach. He claims that its "leading-edge" facilities will allow "brute force" sequencing leading to the rapid analysis of regions of the genome of interest.

Cohen adds that his departure from CEPH will put an end to the controversies that have dogged him over alleged conflicts of interests between his position at the

centre and his involvement in private biotechnology companies (see *Nature* 368, 1775; 1994). "Now I am private," he says.

But Jean Dausset, the founder and president of the centre, describes Cohen's departure as a major loss and the "end of an era". Cohen has been director of the centre since it was created in 1982 by Dausset, the 1980 Nobel joint prizewinner in medicine or physiology, to carry out genetic mapping for the study of multigenic diseases.

Indeed, CEPH has since made an important contribution to genetics by making DNA from a large group of families freely available to laboratories worldwide. Later, in a joint venture with the French Muscular Dystrophy Association (AFM), it created the G  n  thon laboratory, whose successes include physical and genetic maps of the entire human genome.

CEPH will now abandon genetic mapping, but will continue to use maps to identify genes for disease susceptibility. Dausset says that relaunching CEPH's activities will depend on finding a geneticist of international renown to lead the centre. But he accepts that Cohen was "right to leave", given that CEPH could never provide him with the financial resources needed for his ambitious plans in genome research.

Nevertheless, the separation between Cohen and CEPH has not been without its problems. One issue that remains to be resolved is the fate of the techniques and materials developed by Cohen at CEPH. Dausset says that there is a "real risk" that Cohen will take most of CEPH with him when he leaves. Both men are negotiating the extent of transfers of staff and resources.

**Declan Butler**

# UK chemists seek a greater role in policing arms ban

**London.** British chemists are mounting a last-minute appeal to the government to give academics a greater voice in the implementation of restrictions on research materials which could, in larger quantities, be used to manufacture chemical weapons.

The restrictions are included in the measures that the government is planning to adopt in order to comply with the Chemical Weapons Convention. The text of this treaty was agreed in principle in 1993, and now awaits ratification by at least 65 states before taking effect — which is widely expected to occur within the next 12 months.

As part of its implementation plans, the British government is proposing that all researchers using the chemicals listed in the proposed treaty will have to be registered with the Department of Trade and Industry (DTI), the body that will be responsible for ensuring observance of the treaty. Under certain circumstances, indeed, such research could even be banned.

Last week, during a debate in the House of Lords, the government released further details of how its licensing process will operate. In particular, Lord Fraser, minister of state at the DTI, specified that scientists who use small quantities of agents such as cholinesterase inhibitors for studies of nerve

Queen Mary College, University of London, told the House of Lords.

One argument supported by Peston and others was that the best way of avoiding this danger would be to create a statutory advisory committee to oversee the way in which regulations are applied. Members of the committee would include individuals nominated by both the academic and industrial research communities.

"Everyone knows that the government is going to need advice on controlling potentially dangerous chemicals, and we feel strongly that it would be a good idea if the mechanism for this were to be identified in a formal way [in its bill]," says Stephen Benn, parliamentary affairs officer of the Royal Society of Chemistry (RSC). "The public interest is best served by transparency."

Benn says that the RSC is particularly concerned, for example, that despite the considerable powers to be given to the Secretary of State for Trade and Industry under the bill to regulate research, no details have yet been given of any appeals procedure for scientists whose applications to use particular chemicals are rejected.

The government has so far resisted pressure to set up such a statutory advisory group. In last week's debate, Fraser repeated previous arguments that such a committee might result in extra bureaucracy and reduce the flexibility involved in obtaining advice through non-statutory channels.

Such arguments are supported by the chemical industry, which also claims that an advisory committee could raise problems of confidentiality. "The wider information goes, the greater the risk of losing control of confidentiality," says David Culpin of the Chemicals Industry Association.

But Alistair Hay, senior lecturer in chemical pathology at the University of Leeds, warns that, without a panel of outside advisers, the government could end up listening primarily to its own chemical warfare experts and representatives of the chemical industry.

"I do not see that there is anyone [in the government's plans] who, while being sympathetic to the need to control chemicals, would say to the government: do you realize what this might mean for the academic community up and down the country," says Hay.

The proposal for an advisory committee received considerable bipartisan support during last week's debate when several of those present expressed concern that, under current plans, the task of monitoring and verifying compliance with the proposed convention would be carried out by only seven DTI officials.

**David Dickson**

# Legal action 'may end search for truth' in Canada's HIV inquiry

**Toronto.** Canada's long-running inquiry into blood products contaminated with HIV and hepatitis has taken a new turn: more than 40 organizations and individuals whose activities have come under scrutiny have decided to take legal action to prevent the inquiry from concluding that there were individual cases of misconduct.

Their action follows advance warning by Mr Justice Horace Krever, head of the inquiry, that he might reach such a conclusion in his report, intended to give those likely to be named a chance to respond.

But the groups, which include the Canadian Red Cross, the federal government, pharmaceutical companies and every province except Saskatchewan, have filed suits in Canada's federal court asserting that such a course of action is beyond the inquiry's mandate.

The lawsuits have attracted criticism from organizations representing individuals affected by transfusions of infected blood. The president of the Canadian Haemophilia Society, for example, claims that, if the suits are upheld, the truth will never be known about why more than 1,200 haemophiliacs and transfusion recipients contracted HIV between 1980 and 1985.

The Canadian Red Cross in particular faces more than 70 allegations against senior officials. The charges include failing to take basic measures to exclude people in high-risk groups from donating blood, as well as continuing to distribute blood products that had not been heat-treated when it had a substantial inventory of such products.

Among those named is George Weber, the secretary-general of the International Federation of Red Cross and Red Crescent Societies. Weber was head of the Red Cross in Canada between 1983 and 1993, and is alleged to have failed adequately to "oversee, direct and provide resources" for blood transfusion and donor recruitment services.

But in an affidavit aimed at preventing Krever from making such charges part of his report, Weber says the notice of potential findings "could seriously damage my career and future employment prospects". The groups taking legal action also argue that Judge Krever's actions will expose them to claims for compensation from those who contracted HIV or hepatitis C after transfusions with infected blood.

The report of the Krever commission is due to be published by the end of September, but seems certain to be delayed by the legal action. Last Friday, 2 February, consumer groups opposed to the lawsuits were given the right to participate in hearings on the suit, due to take place in May.

**David Spurgeon**



and brain function will be limited to five grams a year for a general licence.

Research chemists emphasize that they strongly support both the proposed treaty and the United Kingdom's plans to ratify it. But some are concerned that an excessively rigid application of restrictions on the use of potentially dangerous chemicals could unnecessarily impede research.

Their views were expressed by several speakers in last week's parliamentary debate. "It is vital that we rid the world of chemical weapons, but we do not want to rid the world of the science of chemistry," Lord Peston, previously professor of economics at



## Biotech shares fall as company halts asthma drug trials

Oxford. Britain's burgeoning biotechnology sector lost some of its gloss last week when Celltech, one of its most well-established companies, announced it was halting trials with CDP840, a potential asthma drug from a class of drugs known as selective phosphodiesterase IV inhibitors (PD IV).

The value of Celltech's shares on the London Stock Exchange dropped by more than 25 per cent immediately following the news, and the slippage then continued. The shares of other British biotechnology stocks fell in sympathy.

Celltech and its partner, Merck Sharp and Dohme, the US pharmaceuticals company, decided to halt the phase II trial, which identifies the effectiveness of a drug against a target disease, because CDP840's performance did not match expectations.

"In collaboration with Merck we identified a number of desirable properties a new asthmatic treatment should have," explains David Bloxham, chief executive of Celltech Therapeutics. "It should be orally active, have anti-inflammatory activity and be able to prevent rapidly the initial contraction of air pipes that causes asthma."

The two companies were hoping to develop a product combining the rapid relief of symptoms associated with beta agonists such as salbutamol with the long-term anti-inflammatory activity of the steroid drugs at present prescribed to asthmatics.

"While the phase II trial demonstrated that CDP840 is orally absorbed, safe and had anti-inflammatory activity, it did not provide the desired rapid onset or relief," Bloxham adds. "To have a new asthma medicine that can be taken orally is not enough in these cost-conscious times."

Celltech's action reflects the desire of most biotechnology companies to axe engaged new products as early as possible in the development programme if they do not meet projected goals; more than 80 per cent of the costs involved in developing a new medicine are associated with phase III clinical trials and beyond.

Indeed, products that get as far as phase II still have only a 50 per cent chance of reaching the market. But when a product is abandoned, the investment community tends to use this to judge a company's performance, as it has little other information on which to base an assessment.

Although Celltech's decision may be the end of the road for the CDP840 molecule, it is not necessarily the end of PD IV as an asthma therapy. Bloxham claims that CDP840 has provided both Celltech and Merck with information that they plan to use in the design of newer PD IV molecules.

Mike Ward

## 'Fraud police' face dilemma over warning medical boards

San Diego. Officials from the US Department of Health and Human Services (HHS) have said they will discuss whether medically qualified researchers who admit to fabricating research data should be automatically reported to the state-run bodies that issue licences to practise medicine.

This follows evidence of at least one case in which such information was not passed on by the National Institutes of Health's Office of Research Integrity (ORI). According to ORI officials, this was partly out of concern that any such ruling would discourage researchers from admitting their guilt, thus increasing the workload on the office, and delaying the consideration of other cases.

The case involves a former research fellow in obstetrics and gynaecology at the University of California at San Francisco (UCSF), James T. Kurtzman, who admitted in 1993 to fabricating data on nitric oxide levels used to measure uterine function.

Following an investigation by officials from UCSF and the National Institutes of Health (NIH), which funded the research, Kurtzman left San Francisco quietly in 1994, eventually agreeing to a penalty under which he was barred for three years from involvement in federally funded research.

The agreement, which was reached last year between Kurtzman and the ORI, specifically states that Kurtzman's admission to falsifying research would not be used to prevent him from practising medicine. But board officials who, as a result, were not made aware of his misconduct are upset.

"We want to know about cases like these," says Candis Cohen, a spokeswoman for the Medical Board of California which, like similar bodies in other states, can impose penalties that can include revoking the licence of physicians found to have engaged in scientific misconduct.

Lyle W. Bivens, director of ORI, acknowledges that state medical boards are keen to learn about cases of scientific misconduct that may raise questions about the integrity of licensed physicians. But he says that, as a rule, his agency does not notify the board. "We wouldn't get voluntary exclusion agreements if we pulled the rug out from under a person's livelihood," says Bivens.

"This is a tough issue," he admits. "It is one we haven't struggled with." Whether someone who admits to having taken short cuts or fabricated research results is inevitably a bad physician "is not a judgement we can or should make," says Bivens. "But it does beg the question of whether medical boards should be notified."

Arthur J. Lawrence, a senior adviser in the office of the assistant secretary of the HHS, which would be responsible for draw-

ing up any federal guidelines on the issue, admits that Kurtzman's cases "poses a real interesting question". He says that HHS attorneys will review the issue.

Some researchers believe that colleagues who admit to fraud should be reported to medical authorities. These include R. Kirk Riemer, who headed the UCSF research project and says that over two years of his work were ruined by Kurtzman. "It's wrong to even consider that this behaviour is not going to carry over to other aspects of life, professional or private," said Riemer.

Insisting that he has since tried to "take the high road", Kurtzman describes the episode as "the most painful thing in my life". He now teaches in a UCSF-affiliated residency programme at a community hospital in Salinas, where he was recently given a teaching honour and a community service award for treating the poor.

But in California, as in many other states, hospitals must report physicians who leave their medical staff while under investigation. Kurtzman said that UCSF did not report him to the medical board. Karl J. Hittelman, an associate vice chancellor at the medical school, says that universities are caught between laws that require reporting of such events and the possibility of being sued for limiting a physician's livelihood.

"Whether intellectual dishonesty at the research bench can be assumed to have an implication for medical practice is a complex issue that is saturated with legal ramifications," said Hittelman. "At a personal level, it sends up a flashing yellow light in my brain. But it needs to be addressed on a case-by-case basis."

Medical board officials agree — but say that it is precisely for this reason that they are keen to be notified of such cases. California is already considering what to do in the case of Anand Tewari, who was barred in 1994 from receiving federal research funds for five years after he was found to have fabricated research results while a fellow in surgery at Stanford.

According to federal records, Tewari signed an agreement with ORI similar to Kurtzman's, in which he acknowledged faking laboratory results on a number of patients in a drug study, but both sides agreed that this should not affect his ability to practise medicine.

In this case, however, Stanford notified the medical board, which in September filed an challenge to Tewari's medical licence, alleging dishonesty and gross negligence for the research misconduct. A state hearing is to be held to decide whether Tewari, who could not be reached for comment, should be punished in any way.

Rex Dalton



# Congress urged to end budget uncertainty

**Washington.** Ninety US congressmen have signed a letter to the leading members of a budget committee in the House of Representatives calling for immediate agreement on a budget up to the end of the financial year for the National Science Foundation (NSF), the agency that funds most non-biomedical university science.

At present, the foundation has had its funding approved only up to 15 March, a fallout from disagreements between Congress and the Clinton administration. The letter calls on Robert Livingston (Republican, Louisiana), chair of the House appropriations committee, and Jerry Lewis (Republican, Missouri), chair of its subcommittee with jurisdiction over NSF, to "do whatever is possible to provide full-year funding for the NSF as soon as possible".

The letter was circulated last week by Vernon Ehlers (Republican, Michigan), a former research physicist and increasingly prominent defender of science on the House Science committee. It follows an

orchestrated barrage of faxes to Congress from scientists in various disciplines drawing attention to NSF's plight.

Several scientific societies tried to rally support for NSF at the end of last month, following a speech from Neal Lane, the director of NSF, in which he criticized "the perceived stony silence" of the research community about the plight of the agency. At that time the NSF had no budget agreed for the 1996 financial year and faced the prospect of yet another shutdown (see *Nature* 379, 281; 1996).

The immediate threat to NSF was subsequently lifted, but only until March. The scientific societies hope, however, that they have succeeded in raising the profile of the agency, increasing the chances that the NSF will get full-year funding after that, and even improving its funding prospect for next year.

The American Physical Society (APS) contacted all its members by electronic mail two weeks ago, asking them to fax local senators and congressmen with an appeal for

funding for NSF. At least 2,000 communications with Congress resulted from this appeal, says Robert Park of the APS.

In addition, a coalition of biologists — the American Society for Biochemistry and Molecular Biology, the American Society for Cell Biology (ASCB), the Biophysics Society, the Genetics Society and the American Association of Anatomists — asked the 2,000 members of its "legislative network" to contact the congressional leadership. According to James Bernstein, director of public affairs at the ASCB, more than half responded. "There clearly is a great deal of support for NSF, even from people who are not NSF grantees," says Bernstein.

There is no immediate prospect of the NSF following the National Institutes of Health (NIH) in obtaining a special appropriation that will assure its funding for the rest of the year. But the scientific societies believe that the exercise could mark the beginning of a more effective lobby for the NSF.

**Colin Macilwain**

## Scientific caution 'blunts efforts' to conserve fish stocks

**London.** Scientists are partly to blame for the perilous state of the world's fish stocks because their cautiously worded advice has been used by fisheries managers to continue the practice of overfishing, thus endangering the long-term sustainability of stocks, according to a report\* published last week by Britain's House of Lords.

The report, prepared by a subcommittee of the Lords' Select Committee on Science and Technology, describes the world's fish stocks as being "in a state of crisis". It calls for broad-based fisheries management panels to help achieve consensus on conserving fish stock and the creation of an International Panel on the Oceans "to help push issues of global concern up the political agenda".

Almost 70 per cent of species are either depleted or being fished at around the maximum sustainable yield, the report says. Yet national catch quotas remain high because of concern about the political implications of unemployment in the fishing industry.

The report says that unless scientists deliver "clear and forceful" advice, governments will continue to sacrifice the long-term sustainability of fish stocks in favour of protecting "short-term" jobs in the fishing industry. "Scientists will always want to draw attention to any degree of uncertainty in their assessments," it says. "But scientific professionalism is currently providing an excuse for political compromise."

The 77-page report also calls for better cooperation between scientists, fishermen,

fisheries managers and politicians, more technical measures to improve fish stocks and more aquaculture.

The subcommittee heard evidence from a wide range of witnesses. Some scientists volunteered the view that ambiguous scientific advice may contribute to overfishing, pointing out that the science of fisheries management is not yet firm enough to provide reliable data. Unresolved issues include

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

**Fish wars: over-fishing has contributed to conflicts between fishermen, including slashed nets**

uncertainties about natural variation in fish stocks, including natural mortality, and establishing the correct age in a fish's life when harvesting becomes sustainable.

Organizations such as English Nature and the Joint Nature Conservation Committee told the inquiry that higher quality data are needed if scientific advice is to be relied upon. But others disagreed. Michael J. Holden, former head of the Conservation Unit at the European Commission's Directorate General of Fisheries, who died

recently, said politicians are aware that "the scientific basis for formulating fish policy has been available for 60 years". Representatives of Scottish National Heritage also claimed that current scientific methods to provide reliable advice for fish stock management are adequate.

John Shepherd, director of the new Southampton Oceanography Centre, said that the demands being imposed on science are more than it can bear. But he conceded that the International Council for the Exploration of the Sea, the body responsible for carrying out scientific assessments of fish stocks, should "revert to blunter and bolder advice rather than the rather heavily qualified advice" it produces at present.

Lord Butterfield, a member of the subcommittee and a former vice-chancellor of the University of Nottingham, said the committee recognized that "unresolved issues in fisheries science" made effective fisheries management more difficult. The committee, he says, also recognized that scientists were constrained by the nature of their training from making value judgements.

"I do understand their dilemmas," says Butterfield. "But who is then going to fill the gap between forming the data and forming opinions?" He adds that scientists "should say clearly that there are certain visible trends, which, if allowed to continue, will lead to catastrophe. I feel if they had said this earlier, it would have had an impact on policy-makers."

**Ehsan Masood**

\*Fish Stock Conservation and Management, Second Report (Session 1995-96), ISBN 0 10 403 0968, HMSO £12.20.

## UK defence ministry plans renewed assault on Gulf War syndrome

**London.** Britain's Ministry of Defence has ordered a fresh study into claims that a group of armed forces personnel are suffering from Gulf War syndrome, a mysterious illness related to their service during the 1991 Gulf War.

Fifty-one-thousand British service men and women went to the Gulf. Seven hundred veterans have since complained of a range of illnesses including memory loss, skin complaints and chronic fatigue. Others have cited birth defects in 60 children born after the Gulf War as evidence that Gulf War syndrome exists.

A three-year epidemiological study, overseen by the Medical Research Council, will be commissioned to compare the health of Gulf veterans with a similar group of personnel who did not serve in the Gulf. A similar study will try to establish whether the defects in children born after the war is statistically significant.

An earlier examination of 350 veterans by the ministry found no evidence of a common syndrome, though some were found to be suffering from serious illnesses. The veterans initially suggested that the disorders may have been caused by chemical or biological attacks during the war. Another theory suggests Gulf War syndrome may be the side effects from the mixture of vaccinations given to protect against chemical and biological attacks. □

## US-Russian projects announced

**Washington.** A joint oceanographic survey by the US and Russian navies, to be conducted in the Sea of Okhotsk in eastern Russia, and increased US government backing for a research fund to support scientists in the former Soviet Union, were two results to emerge from last week's sixth meeting of the Gore-Chernomyrdin Commission to promote economic and technological cooperation.

The National Science Foundation (NSF) announced that it will

contribute an additional sum, as yet unspecified, to the \$10-million Civilian Research and Development Fund (CRDF) established last year with matching funds from the US government and philanthropist George Soros. The NSF contribution will pay the expenses of US scientists in collaborative projects who are limited to receiving 20 per cent of the funds awarded under current CRDF grants.

The National Institutes of Health (NIH) announced its own \$1-million contribution to the CRDF, to be matched by another \$1 million from a US defence conversion fund. The NIH is particularly keen to collaborate with former Soviet scientists who worked on biological and chemical warfare, and may be able to provide insights into emerging infectious diseases. □

## Cell biologists celebrate prize

**London.** Three cell biologists have been honoured in the 1996 King Faisal International Prize awards for science and will share the \$200,000 prize. The winners, who each also receive a gold medal, are: Hugh Pelham, Head of the division of Cell Biology at the UK Medical Research Council's Laboratory of Molecular Biology in Cambridge; Günter Blobel, John D. Rockefeller, Jr Professor at the Laboratory of Cell Biology, Howard Hughes Medical Institute, Rockefeller University, New York; and James E. Rothman, Program Chairman in Cellular Biochemistry and Biophysics at the Memorial Sloan-Kettering Cancer Center, New York, USA. □

## Top-up funds for EU research

**Munich.** The European Commission has put forward a formal proposal to the European Parliament and Council of Ministers for using 'top-up' funds for the fourth Framework Programme. It proposed that the funds should be used to support five task forces that were created by the commissioners for research, industry and transport last year to coordinate and strengthen European Union research activities in key areas, particularly those relating to transport.

In principle, an additional ECU700 million (US\$889 million) is

# Did you know Pharmacia Biotech has spent over 23 years perfecting the technique that purifies mRNA?

Sure, you can purify an mRNA sample in under an hour. But as mRNA is one of the most delicate molecules to purify, you shouldn't take any unnecessary risks. Why not demand more than just speed out of your mRNA purification products?

Look for experience. We've been developing the technique used to purify mRNA, affinity chromatography, for over 23 years. That gives us the experience and knowledge to develop products for mRNA purification that are fast, safe and effective.

So what are our products for mRNA purification like? Today, QuickPrep® Kits not only let you purify mRNA in under

an hour—they also generate up to 5.4 µg of polyadenylated RNA from 100 mg of starting tissue. What's more, QuickPrep Kits let you purify mRNA directly from sample sizes as small as a single cell.

The next time you're about to purify mRNA, put products developed through extensive experience to work for you. Here's how to do it. Just call us at 1 (800) 526 3593 in the United States or +46 18 16 5011 from the rest of the world. They're the phone numbers linked to 23 years of mRNA-related purification experience.

available for EU research programmes beyond that already committed to Framework projects, provided that the council of ministers and parliament agree to release the money. The council, which includes many dissenters, will respond to the commission in March, and a final decision could be made by the summer. □

## Japan tracks infected blood supply

**Tokyo.** A decade after thousands of people in Japan were infected with HIV through contaminated blood products, the head of the Ministry of Health and Welfare, Naoto Kan, last week ordered pharmaceutical companies to reveal which medical institutions received blood products that may have been contaminated with the virus.

Several thousand haemophiliacs are thought to have been infected with contaminated blood coagulants in the early to mid 1980s. But the minister's directive is primarily aimed at tracing non-haemophiliacs who have been infected with HIV. An earlier request to the companies revealed that at least 322 non-haemophiliacs received tainted products, and 11 of them are HIV positive. A further 274 have yet to be tested.

Until now, attention has focused on the infection of haemophiliacs, and Japan has only recently become aware that many non-haemophiliacs may also have been infected with blood products that were not heat-treated to kill HIV. □

## US 'will honour' uranium pledge

**Munich.** Hazel O'Leary, US Secretary of Energy, has announced that the United States will honour an agreement made in 1978 to take back spent nuclear fuel from foreign research reactors that have agreed to convert from highly-enriched uranium to low enriched uranium.

According to O'Leary, the move, which means that the United States is now ready to take back 20 tonnes of weapons-grade uranium from about two dozen European research reactors, is part of the US commitment to preventing nuclear weapons proliferation.

The agreement had come under attack from environmental lobbyists, particularly in Savannah River in South Carolina, where the interim storage of this fuel had been planned. All shipments of spent fuel had been suspended, pending a full environmental impact statement. This listed all global options for disposing the waste, and was presented to the Department of Energy (DoE) at the end of last year, enabling O'Leary to overcome the legal challenges that had been filed against the DoE in South Carolina. □

## Ocean research sounds fine

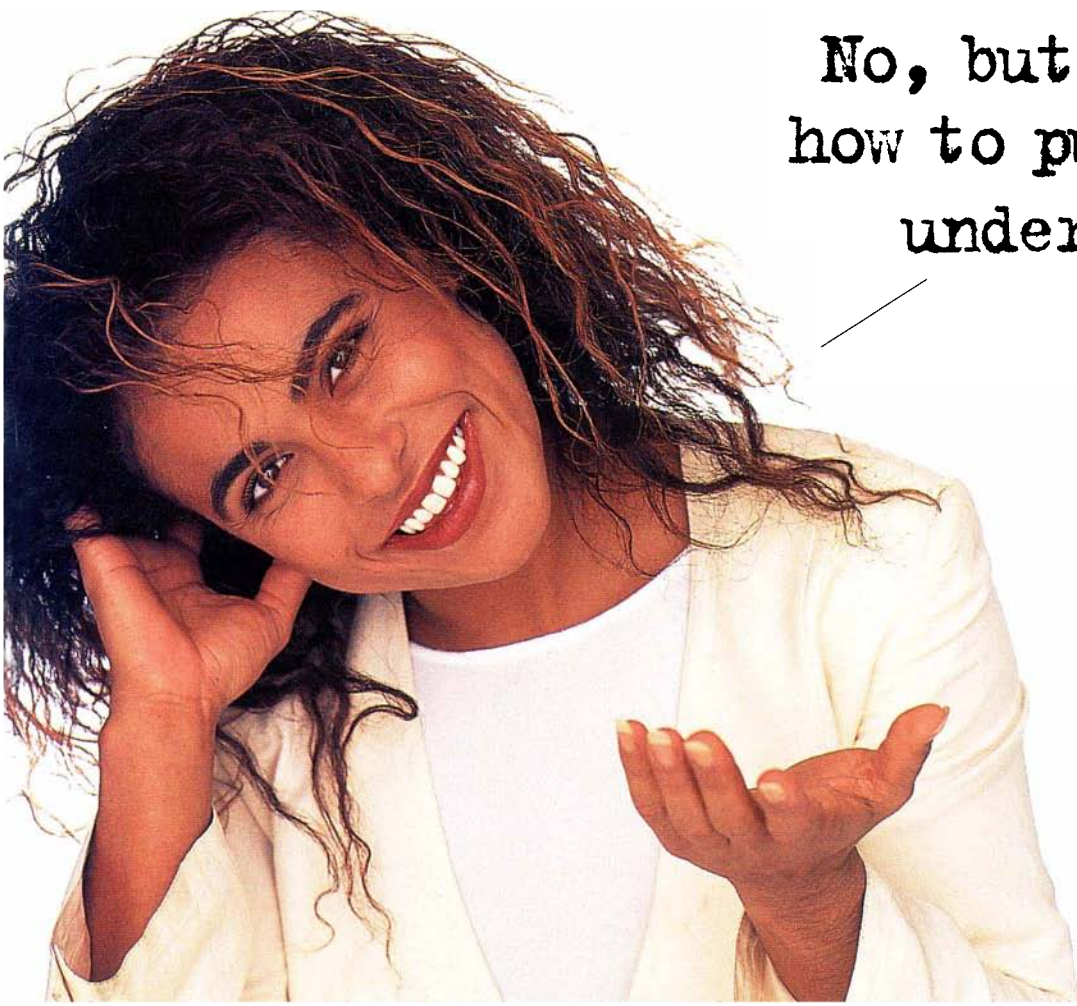
**Santa Cruz, California.** The techniques planned for the Acoustic Thermometry of Ocean Climate experiment, designed to study global warming by producing sound waves in the Pacific off the California Coast, appear to have no damaging effects on marine mammals.

In the experiment, loud-speakers distributed in 1,000 metres of water send out 20-minute, 60-90 Hertz coded signals of up to 195 decibels to measure ocean temperature. Environmentalists feared that the sound would harm the animals, especially sperm whales which can dive deep enough to be near the undersea speakers.

But Dan Costa, a biologist at the University of California at Santa Cruz, said that after several tests biologists could discern no changes in the behaviour or status of whales and dolphins. □

## Travel costs under fire at NOAA

**Washington.** Staff at the US National Oceanographic and Atmospheric Agency (NOAA) are under fire from Republicans in Congress for their high travel costs and alleged misuse of government-issued travel charge cards. Robert Walker (Republican, Pennsylvania) says he is "deeply concerned" about the science agency's \$35 million-a-year travel costs, and has demanded, among other things, a list of the names and job descriptions of the 5,000 NOAA staff issued with the American Express travel cards. A spokeswoman for NOAA pointed out each staff member is responsible for paying off the balance on their own card. □



No, but I know  
how to purify it in  
under an hour



**Pharmacia  
Biotech**

Uppsala, Sweden. (And the rest of the world)



# Why genes can be patented

SIR — Fergus Davison highlights a difficult point in the understanding of patents derived from genetic studies (*Nature* 379, 111; 1996). I, too, thought George Poste (*Nature* 378, 534–536; 1995) skipped over the distinction between discoveries and inventions.

Perhaps some enlightenment can be given by reference to another well known example. Antibiotics have been patented for years without the challenges that have been applied to DNA sequences, but they are also natural molecules produced by living organisms and are, in a very real sense, discovered. Why, then, are they classed as inventions and therefore patentable?

The fact is that the antibiotics produced by living organisms in their wild state are not patentable. However, if an individual or a company searches for and finds the right organism, identifies a useful property of one of its metabolites, isolates, purifies and characterizes the product and devises a method of making and using it, then the outcome is not the antibiotic in its natural state but the genuine product of human ingenuity. It is therefore classed as an invention. The route to finding new antibiotics was used many times over, but the novel chemicals that emerged remained patentable. It is greatly to our advantage that this has been and continues to be the case. Discovering and developing new antibiotics, as with all drugs, is an expensive process, and the patent system is a major mechanism whereby society can encourage scientists in or outside companies to search for new remedies for the general good.

The same arguments apply to DNA sequences. In their natural state, within the body of an organism, directing the production of enzymes, hormones or whatever, they are not patentable. But a DNA sequence that can specify erythropoietin, for example, when identified, characterized and transferred to a suitable production, can give a product that, when purified and properly formulated, is of great value to some patients. In this process, neither erythropoietin nor its gene is in its natural state, and the DNA sequence, as a component of the system that can make such a useful product, is justifiably patentable. Some of the techniques might seem routine to an expert in gene cloning, as would some of the steps in identifying an antibiotic to a microbial ecologist, but the tasks that face Poste and his colleagues are far from trivial, and if society wants new medicines it had better make sure that someone is encouraged to discover them and make them available.

On the arguments given above, Newton could not have patented gravity any more than someone who had just discovered its

secrets could patent the electrical phenomena that comprise the lightning of an electrical storm. However, if someone could take components of those phenomena and reproduce them in a laboratory in a controlled way and then show how they could be used, should there not be a mechanism for ensuring that that person is recompensed for the benefit that accrues to society? Those who have difficulties with DNA patents should consider that question each time they turn on an electric light.

**Norman H. Carey**

*Russells Close, High Street,  
Chinnor, Oxfordshire OX9 4DJ, UK*

## Chance remarks

SIR — Every week, millions of British people buy national lottery tickets, all hoping to be the one who accurately predicts 6 numbers lying between 1 and 50. Every other week or so, someone succeeds. The winner is not clairvoyant, merely living proof that if enough trials are performed, even extraordinarily unlikely events may be observed. Along similar lines, most scientists know that a table containing 20 independent statistical tests for a non-existent trend will, on average, yield one value indicating significance with  $P < 0.05$ . Such a value is referred to as a 'type I error', and should properly be ignored.

However, it is interesting to go beyond a single paper to look at how type I errors might operate between laboratories. Consider the fictitious fate of 20 behavioural ecologists who encounter the same provocatively stimulating dataset at a seminar given by a colleague. All 20 privately develop the same exciting hunch which they then test independently, each with a single experiment. Even if the common hunch has no basis, the odds favour one of these scientists finding a 'significant' value at the  $P < 0.05$  level, which might get published.

Following publication, 1,000 desperate lecturers searching for interesting undergraduate projects all opt for a simple one-experiment test of this new finding. The projects are completed and 50 yield significant values at the 0.05 level, 10 reach 0.01 and one even attains 0.001. With a bird's-eye view, a neutral observer would realise that there all these 'significant' values are likely to be type I errors. However, from the point of view of the individual, there are now 50 excited supervisors of whom perhaps 30 prepare for immediate publication.

Now, suppose the remaining 20 lecturers are more cautious and tell their students to repeat the experiment. One person manages to reproduce his or her results and publishes a convincing paper, whilst the others shake

their heads and ascribe their fluctuating fortunes to student error. Of course, the bird's-eye observer is still not surprised: there are 20 new tests and one yields a  $P < 0.05$ . Of course, the happy 'winner' is blissfully unaware that the result is a type I error, and the scientific community, faced with a widely publicized seminal paper followed swiftly by 31 pieces of supporting evidence, one of which appears really quite convincing, now believes the trend to be real. Yet it never existed.

Finally, we should bear in mind that experiments do not have to be related. Most statistical tests are essentially equivalent, and ask simply whether there has been a significant deviation from random. All over the world, vast numbers of such tests are performed daily, spread across many disciplines. Where trends exist, they will tend to be detected. However, among the vast number of tests applied to nonexistent trends, many spuriously significant  $P$ -values will arise, some of which will be extreme. The exact number of artefactual findings will be related to the number of tests that 'fail', which is never likely to be determined. Consequently, it is perhaps worth bearing in mind that there is such a thing as a global type I error, and that individually good statistical practice does not necessarily remove the risk.

**Bill Amos**

*University of Cambridge,  
Department of Genetics,  
Downing Street,  
Cambridge CB2 3EH, UK*

## Well reviewed

SIR — As one of the lead authors of the Intergovernmental Panel on Climate Change (IPCC) 1994 Scientific Assessment, I must protest about Richard Courtney's ridiculous claim (*Nature* 379, 109; 1995) that "IPCC approved the summary of its 1994 Scientific Assessment before that report was written".

The summary of the 1994 report was approved in Maastricht in September 1994. By that time, the report was virtually complete, having undergone, first, peer review in February–April 1994 and, after modification, a review by each country during May–July 1994.

This review process was not a minor one — it entailed hundreds of copies of the drafts being sent out for review. Indeed, I would be very surprised to learn, given Mr Courtney's interest in the IPCC process, that he had not seen at least one of these earlier drafts.

**Keith P. Shine**

*Department of Meteorology,  
University of Reading,  
2 Earley Gate,  
Reading RG6 2AU, UK*

# Visual perception: rivalry and consciousness

Francis Crick

**As a paper elsewhere in this issue attests, the problem of the neural basis of consciousness looks ever more tractable as neurobiologists delve into the process of visual perception.**

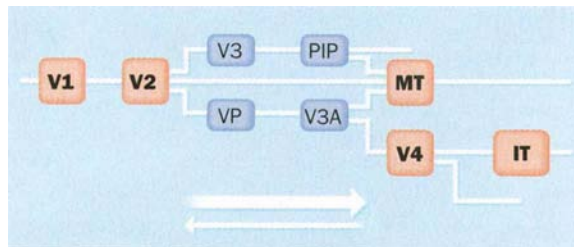
WHAT do you see if, by some optical trick, your right eye is shown something quite different from your left eye? One might reasonably guess that you would see both scenes, one on top of the other. For a very brief period, this may happen; but after that, the two percepts alternate at irregular intervals, changing every few seconds. The brain allows you to perceive only one of them at a time. This is called binocular rivalry.

What goes on in your brain during binocular rivalry? It has been widely assumed (see ref. 1 for example) that rivalry happens at an early stage in visual processing, when the images from the two eyes are kept somewhat separate. On the other hand, a new paper by Leopold and Logothetis (page 549 of this issue<sup>2</sup>) suggests that rivalry occurs at several levels in the visual pathway. They came to this conclusion by studying, not humans, but two macaque monkeys, using methods somewhat similar to those used earlier by Logothetis and Schall<sup>3</sup>. The monkey fixated a spot of light (to prevent it from moving its eyes) and was trained to report, by pressing one of two levers, which of the two alternatives it saw. Elaborate precautions (described in the paper) were taken to make sure the monkey did not cheat, as monkeys will do if given half a chance.

Logothetis and Schall<sup>3</sup> took their electrophysiological recordings from the midtemporal cortex (MT), using moving stimuli because the neurons there respond preferentially to motion. Leopold and Logothetis<sup>2</sup> have recorded from near the border of the first and second visual cortical areas (which they refer to as V1/V2) and from cortical area V4 (see figure), the neurons of which respond preferentially to shape and colour. Further experiments on inferotemporal cortex (IT) are in progress<sup>4</sup>. The visual stimuli used in this paper<sup>2</sup> were small patches of bright and dark lines in a particular orientation, chosen to be optimal for that particular neuron under conditions of non-rivalry. The state of binocular rivalry was created by presenting the two eyes with similar patches, but with the lines on them at right angles.

The striking result was that whereas

some of the neurons correlated with what the monkey reported seeing, the firing of the majority remained unaffected by rivalry. Such neurons were found at all three locations in the visual system (V1/V2, V4 and MT). From this, the authors concluded that rivalry cannot be a process occurring, once and for all, at V1 or earlier. Their conclusion is further supported by psychophysical experiments on humans (N. K. Logothetis, D. A. Leopold and D. Sheinberg, personal communication).



Simplified representation of part of the visual pathway, the boxes representing functional subdivisions of the visual cortex. Those mentioned in the text are in red. The main flow of information is from left to right, but all connections go in both directions (many have been omitted — see ref. 10). Visual input from the retina is transmitted through the thalamus to V1, the primary visual cortex. It is subsequently processed by a large number of higher cortical areas (at least 30), of which only a few are shown here. The paper by Leopold and Logothetis<sup>2</sup> discussed here concludes that binocular rivalry occurs at various levels in the pathway, not simply at V1 or earlier. MT, midtemporal cortex; IT, inferotemporal cortex.

Binocular rivalry is a phenomenon of interest in its own right; but its real importance, as John Allman realized some time ago<sup>5</sup>, is that it may shed light on the baffling problem of visual awareness, a visual form of consciousness. In very simple terms, the visual input is relatively constant yet the percept changes radically with each alternation. Which neurons follow the visual input and which follow the percept? The neurons most closely involved in what is actually seen are likely to be some of those whose firing represents what the monkey reports it is seeing.

It is a basic assumption of all such experiments that the visual awareness of the macaque monkey is not unlike our own, even though a monkey cannot make verbal responses. Of course, it is not easy to get the monkey to tell us what it is seeing (though it has been done — see the ingenious paper by Cowey and Stoerig<sup>6</sup>). It could conceivably be responding in a rather automatic way, without any visual

awareness. Leopold and Logothetis<sup>2</sup> make a good case that during binocular rivalry the monkey is very likely to be seeing the alternating images, because all the studied aspects of the psychophysics of the responses are similar to those shown by humans, who certainly can report such percepts verbally. Similar arguments make it unlikely that the monkey was simply responding at random.

The experimental results are not completely straightforward — the reader should consult the paper itself for the details. The neurons do not divide neatly into two classes: those that fire but ignore the percept and those that respond in the expected way to the percept the monkey is signalling. In the cortical areas described so far, a few neurons follow the percept very strongly, but others follow it to varying degrees. The behaviour of some neurons is distinctly odd. For example, during rivalry, some neurons in V4 and MT fire 'out of phase'; that is, they fire when the percept is at right angles to the orientation they respond to when rivalry is absent.

Which type of neuron tends to follow the percept? Each area of the cortex has a layered structure, and Logothetis now reports<sup>2</sup> that, in his earlier work<sup>3</sup> on MT, post mortem studies showed that all the neurons modulated by the percept were located in the deeper layers (5 and 6) — though not all the neurons in the deep layers were modulated. (The location of the modulating neurons in the new studies<sup>2</sup> can only be obtained post mortem.) This hints that special types of neurons may be involved. The key question is: which part of the

1. Blake, R. *Psychol. Rev.* **96**, 145–167 (1989).
2. Leopold, D. A. & Logothetis, N. K. *Nature* **379**, 549–553 (1996).
3. Logothetis, N. & Schall, J. *Science* **245**, 761–763 (1989).
4. Sheinberg, D., Leopold, D. A. & Logothetis, N. K. *Soc. Neurosci. Abstr.* **22**, 19 (1995).
5. Myerson, J., Miezin, F. & Allman, J. *Behav. Analysis Lett.* **1**, 149–159 (1981).
6. Cowey, A. & Stoerig, P. *Nature* **373**, 247–249 (1995).
7. Weiskrantz, L. *Blindsight: A Case Study and Implications* (Clarendon, Oxford, 1986).
8. Milner, A. D. & Goodale, M. A. *The Visual Brain in Action* (Oxford Univ. Press, 1995).
9. Kolb, F. C. & Braun, J. *Nature* **377**, 336–338 (1995).
10. Felleman, D. J. & Van Essen, D. C. *Cerebral Cortex* **1**, 1–47 (1991).

monkey's brain do these modulating neurons connect with? This will not be easy to discover, but it is essential to do so if we are to understand exactly what is going on during visual awareness. What type, and subtype, of neurons are they? And do they have any unusual properties or connections?

Although a casual reader might not realize it, these two papers<sup>2,3</sup> are among the opening salvoes of a concerted attack on the baffling problem of consciousness. I believe that before confronting all its aspects, we first need to discover the neural correlates of consciousness (often called the NCC). For this task the primate visual system seems especially attractive. The two papers show one direct approach

to the NCC. Another, less direct, is to study cases — loosely lumped together as 'blindsight' — in which, by contrast, a visual input leads to an appropriate motor output without any visual awareness<sup>6-9</sup>. No longer need one spend time attempting to understand the far-fetched speculations of physicists, nor endure the tedium of philosophers perpetually disagreeing with each other. Consciousness is now largely a scientific problem. It is not impossible that, with a little luck, we may glimpse the outline of the solution before the end of the century. □

*Francis Crick is at The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.*

## OZONE DEPLETION

# There's safety in numbers

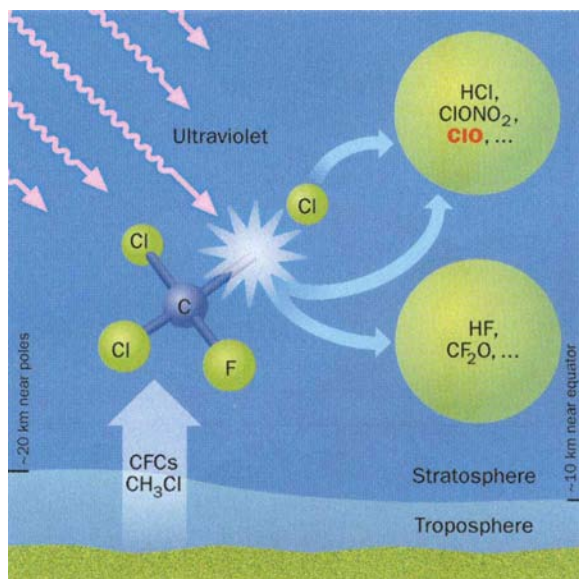
William Brune

SOMETIMES it is important to strengthen an already powerful case. The satellite observations reported by Russell *et al.* on page 526 of this issue<sup>1</sup> add to the weight of scientific evidence that links use of chlorofluorocarbons (CFCs) to stratospheric ozone depletion and that forms the basis of the Montreal Protocol, an international treaty to phase out and control CFCs and other halogen-containing chemicals. Their work also confirms the ideas of Sherwood Rowland, Mario Molina and Paul Crutzen, recipients of the 1995 Nobel Prize in Chemistry. Yet the popular press and some scientists unfamiliar with the evidence remain confused. Some report that the evidence is solid, while others claim that it is the imaginings of environmental extremists.

The scientific case linking CFCs to stratospheric ozone depletion rests on the establishment of three facts. First, that CFCs get into the stratosphere where they are destroyed, becoming the major source of stratospheric chlorine. Second, that chlorine destroys ozone, and by well understood mechanisms. Third, that the amount of chlorine in the stratosphere is large enough to cause the observed ozone depletion. Each of these links has been examined, and borne out, in thousands of studies involving laboratory measurements, atmospheric observations and model calculations<sup>2</sup>.

The observations of hydrogen chloride and hydrogen fluoride made by Russell *et al.* confirm the first link, that CFCs are the dominant source of chlorine in the stratosphere. The data are from HALOE

(Halogen Occultation Experiment) on the Upper Atmospheric Research Satellite, which has been operating since October 1991. The authors choose observations at the 55 km level, where model calculations show that 90 per cent of the chlorine is in



Man-made CFCs and natural methyl chloride ( $\text{CH}_3\text{Cl}$ ) are produced at the Earth's surface, enter the stratosphere and are broken down by UV radiation and subsequent photochemistry. Most chlorine and fluorine becomes HCl and HF respectively. In the Antarctic ozone hole, most of the chlorine is converted to ozone-destroying ClO.

the form of HCl and 80 per cent of the fluorine is in HF. Then, from HCl and HF measurements, they find the increases in stratospheric chlorine and fluorine between 1991 and 1995. These are in excellent agreement with the average increases in total chlorine and fluorine concentration in CFCs and other compounds measured at over 20 ground sta-

tions worldwide between 1985 and 1992 (ref. 2). The growth in CFC emissions was fairly constant during the 1980s, only beginning to fall off in 1992.

To strengthen their argument, the authors compare the absolute chlorine concentration at the surface with that at 55 km. They look at the increases in total fluorine concentration in the stratosphere and at the surface over time, and determine that the average time for CFCs to get to 55 km is 5.9 years. By applying this time lag to anthropogenic chlorine-containing compounds and natural methyl chloride produced at the surface (but not to surface HCl, because it is too water soluble to reach the stratosphere in any quantity), they predict a chlorine concentration at 55 km of  $3.44 \pm 0.23$  parts per billion by volume (p.p.b.v.), only a few per cent different from the observed value of  $3.3 \pm 0.33$  p.p.b.v.

Thus, both the observed increases and absolute values confirm that CFCs and other man made compounds (including carbon tetrachloride, methyl chloroform and HCFCs) are the source of about 80 per cent of stratospheric chlorine. HCl from volcanoes and sea spray can contribute at most only 5 per cent, in agreement with previous observations<sup>3</sup>.

Scientific evidence also supports the assertion that the mechanisms of ozone destruction by chlorine are understood, and that the amount of chlorine is large enough to cause the observed ozone change<sup>2</sup>. HCl is the main component of stratospheric chlorine, but another product of CFC photochemistry, chlorine monoxide (ClO), is the main ozone destroyer. In middle latitudes, the heights at which ClO has its largest effect on ozone depletion are also those where ozone decreases are greatest—below 25 km and near 40 km. That the observed ozone loss is greater than the calculated loss indicates that the understanding of the mechanisms is flawed, but that other processes are not yet included in the models.

The strongest support for these links comes from the dramatic loss associated with the Antarctic ozone hole and the 15–20 per cent loss seen in the Arctic winter stratosphere<sup>4</sup>. These cold polar regions have clouds that initiate a powerful chlorine-dominated chemistry. Aircraft, balloons, satellites and instruments on the ground have all recorded large amounts of ClO and a rapid decline in ozone. The ozone loss matches that calculated by using the observed amount of ClO and other reactive compounds<sup>5-7</sup>.

If the case is already strong, why is this paper<sup>1</sup> an important and timely addition?



Additional evidence is important because a few vocal sceptics are working to undo the Montreal Protocol, a treaty that is already slowing the growth in stratospheric chlorine<sup>6</sup>. These few ignore the weight of scientific evidence, parrot misconceptions that have been thoroughly scrutinized and discredited, focus attention only on the few studies that support their position, and predict that the Montreal Proto-

col will cause economic distress and a worldwide reduction in living standards. By their statements, they have confused the press, the public and members of government.

Deciding whether CFCs affect stratospheric ozone should not be a matter of personal prejudice, but of examining the evidence and drawing conclusions. The conclusion of scientists who have made

the effort honestly to study the evidence is that the use of CFCs has resulted in the loss of stratospheric ozone — certainly in the Antarctic ozone hole and the Arctic wintertime stratosphere, and probably over middle latitudes as well. As Russell *et al.* have said, "Any attempts to conclude otherwise must explain away the scientific evidence provided in this paper". Those who do not believe must provide alternative explanations if future discussions are to be based on science, not politics. □

William Brune is in the Department of Meteorology, Pennsylvania State University, University Park, Pennsylvania 16802, USA.

1. Russell, J. M., Luo, M., Cicerone, R. J. & Deaver, L. E. *Nature* **379**, 526–529 (1996).
2. WMO, *Scientific Assessment of Ozone Depletion: 1994*, Report No. 37, World Meteorological Organization Global Ozone Research and Monitoring Project (1995).
3. Gunson, M. R. *et al. Geophys. Res. Lett.* **21**, 2223–2226 (1994).

4. Waters, J. W. *et al. Nature* **362**, 597–602 (1993).
5. Anderson, J. G., Toohey, D. W. & Brune, W. H. *Science* **251**, 39–46 (1991).
6. Solomon, S. *Nature* **347**, 347–354 (1990).
7. Manney, G. L. *et al. Nature* **370**, 429–434 (1994).
8. Elkins, J. W. *et al. Nature* **364**, 780–783 (1993).

## OBITUARY

## Clair C. Patterson (1922–95)

CLAIR C. (Pat) Patterson died suddenly at his home early in the morning on 5 December 1995. His main contributions to science were the discovery of the age of the Earth and his pioneering studies of global lead contamination. That research, which won him the Tyler Prize for Environmental Achievement, catalysed an unprecedented worldwide policy shift banning lead in gasoline and manufactured products.

Just as Saul Bellow portrayed him in the novel *The Dean's December*, where Patterson appears as Sam Beech, he was justifiably proud of his rural mid-western heritage. Born and raised in Iowa, he earned an AB in chemistry at Grinnell College in 1943, an MS at the University of Iowa in 1944, and a PhD at the University of Chicago in 1951. During that period, he also met two of the people who enabled his scientific career to flourish for half a century: his wife Lorna, whose serenity perfectly complemented Pat's intensity, and his mentor, Harrison Brown, who created an ideal intellectual environment for Patterson to thrive, first at Chicago and later at the California Institute of Technology.

Patterson determined that the Earth and the Solar System are 4.6 billion years old by measuring stable lead isotopes — produced by the radioactive decay of uranium and thorium — in a meteorite, which he assumed had been formed at the same time. This was far older than previous estimates of approximately three billion years accepted at the time. He used the same isotopic techniques to investigate the evolution of the Earth's crust.

Those analyses required the development of 'clean techniques' to eliminate the surprisingly large amounts of contaminant lead in his samples. He traced the contamination to industrial sources, and redirected his research to quantifying the magnitude of anthro-

pogenic lead emissions on a global scale. He found that the emissions began with the smelting of lead ores five thousand years ago, and increased markedly with the addition of lead alkyls to gasoline this century. His analyses of temporal variations of lead

societal costs of environmental lead contamination. However, all of Patterson's research has now been corroborated. His methods have revolutionized environmental and medical research, where clean techniques reveal structure in the biogeochemical cycles of other trace elements, at environmental concentrations often orders of magnitude lower than earlier experiments could have detected.

Patterson was an iconoclast, but as a very private man he disliked being a public figure. At Caltech, he assumed the unique title of Geochemist rather than accept an endowed chair that might interfere with his research. Although he ignored his critics, he was very concerned with the expanding influence of pseudoscientists (whom he classified as degenerative hominids) on future generations of scientists. He was also obsessed with the collegiality, integrity and worth of science, as illustrated in his preface to the volume of *Geochimica et Cosmochimica Acta* (1994) dedicated to him:

True scientific discovery renders the brain incapable, at such moments, of shouting victoriously to the world "Look at what I have done! Now I will reap rewards of recognition and wealth!" Instead such discovery instinctively forces the brain to thunder "WE did it!" in a voice no one else can hear within its sacred but lonely chapel of scientific thought. In that shrine a lasting sense of well-being is conferred through electrifying recognition that the human mind, to which community the scientific mind belongs, does indeed reign supreme over the physical world through its ability to develop understandings of that world's secrets, not to find ways to control it.

Clair Patterson was a scientist for the ages. We will miss him dearly.

A. Russell Flegel

A. Russell Flegel is in the Department of Earth Science, University of California, Santa Cruz, California 95064, USA.

IMAGE  
UNABLE  
FOR A LARGE  
FOR REASON  
REASONS

California Institute of Technology

in ice cores, sediments, oceans, atmospheres and organisms demonstrated that contemporary lead concentrations in the biosphere are often orders of magnitude above the natural levels deduced from comparing the ratio of lead to its biochemical analogue, calcium, in pristine and contaminated tissues. Patterson also put those increases into perspective by showing that lead concentrations in humans are now 500 to 1,000 times pre-industrial levels, and by noting that no threshold concentration for lead toxicity in humans has been established.

His meticulous documentation of lead contamination was initially controversial. Malicious attacks were directed at him by many people because his analyses invalidated their data or their ideas and revealed the enormous

# The ephemeral mountains

Alan Howard

THE Himalayas are being pushed up by tectonic forces and simultaneously eroded down at a fantastic rate. The relationship between uplift, erosion and landscape has remained unclear, but Burbank *et al.*<sup>1</sup> (page 505 of this issue) now show, for the Indus River, that uplift and incision rates are maintained near equilibrium, and that erosion is driven by landsliding from uniformly steep slopes. These processes are probably a characteristic of all young mountain ranges.

Geologists have long recognized that landscapes are created by erosion and deposition acting on tectonically created surfaces, but only during the past few years have geomorphologists and geophysicists become involved in cooperative research<sup>2</sup>. This new impetus to understand the interactions between erosion and deposition has resulted from several developments.

We increasingly recognize that the erosional unloading of mountains and loading from deposited sediment affect the style and rate of tectonic deformation — for example, erosion from mountains lightens the crust, and can cause uplift almost great enough to compensate for the erosion. Our understanding of weathering and erosion has progressed to the point that the development of landscapes can be modelled by computer. Perhaps most importantly, new techniques for quantifying rates of sedimentation, uplift and erosion, and rock ages and exposure ages of land surfaces, allow evaluation of uplift and erosion histories over tens of thousands to millions of years. However, these dating techniques are expensive and time consuming, and they require attention both to the assumptions involved in dating and to determining the history of erosional events that have created the deposits or rock surfaces being dated.

Erosion rates are notoriously difficult to quantify. Measurements based on suspended sediment and dissolved load in rivers inevitably average over large areas and are affected by sediment storage between the source and the measurement site, as well as by climatic and human influences<sup>3</sup>. Recent fluvial terraces can sometimes be radiocarbon dated, and volcanic flows, volcanic ash and loess deposits can be used very locally as datable markers within sediments. In rapidly eroding mountainous terrain, however, rivers may be encased in narrow valley walls with little room for deposits to accumulate. These rivers are often steep and floored with bedrock, so the sediment load is swept through the river system without appreciable deposition.

Such bedrock channels occasionally

shift sideways, either due to meandering or local variations in resistance to erosion. That sometimes leaves isolated portions of the former channel bottom remaining on side slopes as strath terraces. If these abandoned channel bottoms remain essentially unweathered (fluvial scour holes and sediment-polished surfaces remain on the straths studied by Burbank *et al.*), and if appropriate minerals are found in the rocks immediately beneath the surface, the length of time since abandonment of the strath can be found using exposure-age dating.

## IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The valley of the Indus River in the northwestern Himalayas, with a strath terrace (abandoned channel bottom) in the foreground. Extremely rapid incision by rivers and streams in the area, and correspondingly rapid tectonic uplift, leads to a topography determined by landsliding.

This method relies on nuclear reactions induced by cosmic rays in the upper metre or so of the exposed rock<sup>4</sup>. A variety of nuclides accumulate in proportion to exposure time and in inverse exponential proportion to depth beneath the surface. Unweathered grains of the minerals quartz and olivine are particularly suited to this technique. Burbank *et al.* use <sup>10</sup>Be and <sup>26</sup>Al to find strath exposure ages along the Indus River in the Himalayas, and divide the height differences between the strath and the present stream bed by these ages to find average erosion rates.

In igneous rocks, a long-term average rate of erosion can also be estimated from the length of time since mineral grains cooled sufficiently (to below the annealing

temperature) to retain radiogenic isotopes in the crystal structure or the structural damage from their decay (fission tracks). Elapsed ages since annealing can be converted to erosion rates if the geothermal temperature gradient is known and if cooling is driven mainly by exposure due to erosion from above. Ages from apatite fission-track and exposure-age dating agree in the region studied by Burbank *et al.*, although estimated denudation rates are about half as great using zircon, which has a higher annealing temperature — this may indicate more rapid erosion during the past few hundred thousand years.

Estimated stream incision rates vary along the Indus River from 2 to 12 mm yr<sup>-1</sup>. The higher incision rates are correlated with a structural zone that crosses the

path of the Indus and that has probably been undergoing rapid deformation during the past 2 million years.

Because of the difficulty and expense of measuring erosion rates by isotopic dating, geologists would be delighted if the relief characteristics of the landscape could provide an index of erosion rates. If such characteristics could be identified and calibrated with known rate measurements, then, with the wide availability of digital elevation data, erosion rates could be rapidly assessed and mapped. In fact, local relief has been related to denudation rate both in global and regional studies<sup>5</sup>.

Erosional processes are driven by gravity, so one might reasonably assume that more rapid incision would generate steeper relief. Although there is a general correlation between erosion rate and relief in the Himalayas, Burbank *et al.* find little

correlation between slope steepness and incision rate along rapidly eroding sections of the Indus River, despite a sixfold variation in incision rate. They suggest that the steep slope gradients (averaging 32°) in these rapidly eroding regions hover close to the maximum stable angle of the partially weathered bedrock, so that a large increase in downcutting may require only a degree or two increase in

1. Burbank, D. *et al.* *Nature* **379**, 505–510 (1996).
2. Merritts, D. & Ellis, M. (eds) *J. geophys. Res.* **99**, 12135–12315, 13871–14050, 20063–20321 (1994).
3. Summerfield, M. A. & Hulton, N. J. *J. geophys. Res.* **99**, 13871–13883 (1994).
4. Bierman, P. R. *J. geophys. Res.* **99**, 13885–13896 (1994).
5. Schmidt, K. M. & Montgomery, D. R. *Science* **270**, 617–620 (1995).

average slope steepness for landslides to remove material quickly enough. Schmidt and Montgomery<sup>5</sup> have also suggested that slope stability limits maximum relief.

The equilibrium between uplift and erosion is maintained both by the increase in relief accompanying uplift and by increased rainfall over higher ground. Heavier rainfall leads to faster stream incision, and the threshold process of erosion by landsliding carries away material very efficiently.

## GREAT ATTRACTOR

# Meet your cosmic neighbour

Amos Yahil

A MASSIVE cluster of galaxies, lurking closer to us than any other cluster of such a mass, has only now come out of hiding. Although thousands of rich clusters are catalogued, and many have been extensively studied, a nearby example called Abell 3627 has largely escaped attention because it lies close to the Milky Way in the sky. On page 519 of this issue, Kraan-Korteweg *et al.*<sup>1</sup> present convincing evidence that this cluster is comparable in mass to other well-known nearby clusters such as Coma and Perseus. It is located in one of the major large-scale density enhancements in the nearby Universe, known as the Great Attractor, which is responsible for a significant part of the gravitational pull on our Galaxy.

As a fraction of the total mass of the Great Attractor, the cluster's mass is small (around 10 per cent); hence its contribution to the overall gravitational field is less important. Nevertheless it is likely that the cluster points to the bottom of the Great Attractor's gravitational potential well, the position of which is only poorly known from studies of the large-scale distribution of galaxies. This view is bolstered by the recent discovery<sup>2</sup> of strong X-ray emission from the intracluster gas of Abell 3627.

The structure of the Universe within about  $5 \times 10^8$  light years has been extensively mapped in the past two decades, both by directly counting galaxies and by measuring the flows induced by the gravitational effect of density fluctuations. The goals of this research<sup>3,4</sup> are to understand the growth of large-scale structure and to determine the density of the Universe, and so find out whether it is gravitationally bound — so that its expansion will one day reverse — or unbound, fated to expand forever. To make an analogy with geography in the fifteenth and sixteenth centuries, cosmologists are mapping the 'continents' and putting them together to form a unified view of the 'Earth'.

Two hundred years ago, William Herschel already knew of the uneven distribution of galaxies in the sky (although the

We may still hope that other aspects of relief, such as stream gradient or the average spacing between streams (drainage density), depend on stream incision rates in steep terrain. Further development of the approach taken by Burbank and co-workers may well tell us whether that hope is justified. □

Alan Howard is in the Department for Environmental Sciences, University of Virginia, Charlottesville, Virginia 22903, USA.

distinction between external galaxies and gas clouds in our own Galaxy was not yet known), but it was de Vaucouleurs<sup>5</sup>, motivated by a preliminary study of flow velocities by Rubin<sup>6</sup>, who first pointed out a large-scale concentration of galaxies along a great circle on the sky, which he called the Supergalaxy. The high quality surveys of the past decade have shown that most of the large nearby structures are indeed found in and around the supergalactic plane. One of the first of these studies, a survey of the flow velocities of elliptical galaxies<sup>7</sup>, identified a large-scale mass concentration around the Hydra and Centaurus constellations, which became known as the Great Attractor. (The name is actually misleading, implying that it is the main source of the perturbative large-scale gravity we feel. Today we know that other structures and voids exert comparable forces.) Subsequent optical and infrared surveys confirmed the existence of the Great Attractor as a large concentration of galaxies covering about a steradian on the sky, with a density about double the mean density of the Universe (ref. 8 and references therein).

Galaxy and velocity surveys are good at identifying the gross features of large-scale structures but do not generally have the resolution to map them in detail. Even the centres of such amorphous structures are not always obvious, so attention has focused on rich clusters of galaxies within them, particularly those with strong X-ray emission from intracluster gas. The assumption is that the clusters and gas tend to collect at the bottom of gravitational potential wells of large structures. Indeed, rich clusters of galaxies are seen to be the high density 'mountain tops' of larger density-enhancements.

One of the first anomalies noticed about the Great Attractor was the apparent absence of a rich cluster of galaxies at its centre. At first the Centaurus cluster — a strong X-ray source — was suspected to be at the centre, but its galaxies have a bimodal velocity distribution<sup>9</sup> and so it

may be a superposition of two smaller clusters along the line of sight. Recent surveys have also moved the likely centre away from Centaurus.

The centre of the Great Attractor is located not only in the supergalactic plane but also close to where it is intersected by the Milky Way (the two planes are almost perpendicular). This helps to explain why it was not first noticed in surveys, since obscuration by dust in our Galaxy makes the identification of external galaxies more difficult in the 'zone of avoidance' which surrounds the Milky Way. Only in the past few years have deeper surveys attempted to draw back this veil<sup>10</sup>.

The paper by Kraan-Korteweg *et al.* has now identified the central rich cluster at a position only 7 degrees away from the plane of the Milky Way. The cluster had actually already been catalogued<sup>11</sup> with the name Abell 3627, but it was thought to be a poorer cluster because obscuration by the Milky Way reduced the number of galaxies detected, and because the velocity of galaxies within it — a measure of its mass — had been underestimated. The current, deeper search has revealed the obscured galaxies, and measuring the velocity distribution of about 100 of them has established the mass of the cluster to be comparable to that of standard rich clusters of galaxies such as Coma and Perseus.

X-ray emission from Abell 3627 had also gone unnoticed in earlier surveys due to confusion with a source in our own galaxy, but a careful analysis of emission from the cluster detected by the Rosat satellite showed that it has the sixth brightest X-ray flux (in the spectral range 0.1 to 2.4 keV) of all known clusters<sup>2</sup>. Moreover, the mass of the cluster implied by this emission agrees with the estimate based on the velocity dispersion of its member galaxies. It is very likely that the bottom of the potential well of the Great Attractor has indeed been found. □

Amos Yahil is in the Astronomy Program, State University of New York, Stony Brook, New York 11794-2100, USA.

1. Kraan-Korteweg, R. C., Woudt, P. A., Cayatte, V., Fairall, A. P., Balkowski, C. & Henning, P. A. *Nature* **379**, 519–521 (1996).
2. Böhringer, H., Neumann, D. M., Schindler, S. & Kraan-Korteweg, R. C. *Astrophys. J.* (in the press).
3. Dekel, A. A. *Rev. Astr. Astrophys.* **32**, 371–418 (1994).
4. Strauss, M. A. & Willick, J. A. *Phys. Rep.* **261**, 271–431 (1995).
5. de Vaucouleurs, G. *Astr. J.* **58**, 30–32 (1953).
6. Rubin, V. C. *Astr. J.* **56**, 47–48 (1951).
7. Lynden-Bell, D. *et al. Astrophys. J.* **326**, 19–49 (1988).
8. Santiago, B. X., Strauss, M. A., Lahav, O., Davis, M., Dressler, A. & Huchra, J. P. *Astrophys. J.* **446**, 457–471 (1995).
9. Lucey, J. R., Curie, M. J. & Dickens, R. J. *Mon. Not. R. astr. Soc.* **221**, 453–472 (1986).
10. *Unveiling Large-Scale Structure Behind the Milky Way* (eds Balkowski, C. & Kraan-Korteweg, R. C.) (Astronomical Society of the Pacific, San Francisco, 1994).
11. Abell, G. O., Corwin, H. G. & Olwin, R. P. *Astrophys. J. Suppl. Ser.* **70**, 1–138 (1989).



# An elongation factor turn-on

Knud H. Nierhaus

THE elongation factors called EF-Tu and EF-G in bacteria (EF-1 $\alpha$  and EF-2 in other organisms) drive and control protein synthesis on ribosomes by catalysing the elongation of chains of amino acids. Hence their name. They are prototypes of the superfamily of G proteins, which can switch between 'active' and 'inactive' conformations by binding GTP and GDP, respectively<sup>1</sup>. Crystal structures of the various states of EF-Tu<sup>2-5</sup> and that of EF-G with and without GDP<sup>6,7</sup> have told us a great deal about elongation factors and G proteins in general, and also something about ribosomal function. But the mechanism by which EF-Tu switches from the inactive state back to the active one was not clear until now. This gap is closed on page 511 of this issue<sup>8</sup>, where Kawashima *et al.* describe the crystal structure of the complex between EF-Tu and EF-Ts, a protein that mediates the re-activation of EF-Tu.

G proteins are GTPases and regulate a huge variety of cellular processes according to a basic reaction cycle<sup>1</sup>. They adopt an 'on' conformation when bound to GTP. In this state they can bind to another macromolecule or complex and trigger a distinct reaction. After a G protein has done its job the intrinsic GTPase centre hydrolyses the GTP, and the resulting GDP induces an 'off' conformation that leads to the dissociation of the G protein from the macromolecule or complex.

The figure depicts the GTPase cycles of the two elongation factors. The complex to which the factors bind in their 'on' state is the ribosome. In the course of protein synthesis the ribosome oscillates between two main states — that before and that after translocation. Translocation is the relative movement of the complex between transfer RNA and messenger RNA across the ribosome by one codon length: the two tRNAs present in the ribosomal A and P positions move to the P and E positions. The pre- and post-translocational states of the ribosome are separated by high activation-energy barriers (about 90 kJ mol<sup>-1</sup>)<sup>9</sup> which are reduced by the elongation factors, thus enormously accelerating protein synthesis. But these factors also determine the direction of the reaction. EF-G promotes the ribosomal transition from the pre- to the post-translocational state and EF-Tu promotes the reverse reaction. This is why ribosomal systems from all organisms require two elongation factors.

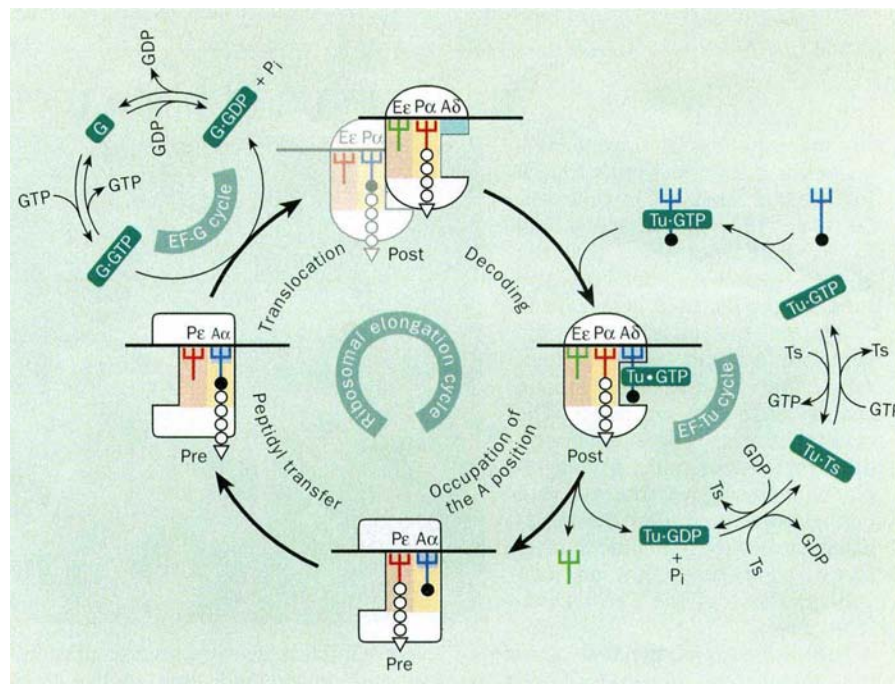
In principle, the GTPase cycle of EF-G is quite simple. EF-G-GTP binds to the pre-translocational state and catalyses the transition to the post-translocational state. The post-translocational state seems to be

energetically favoured over the pre-translocational state; that is, the former can be established without extra energy provided that the activation-energy barrier is kept down. The post-translocational state is able to activate the GTPase centre on EF-G by an unknown mechanism, GTP is then cleaved and EF-G-GDP falls off the ribosome. The intrinsic affinity of EF-G for GDP is low enough to allow its replacement by GTP without other helper molecules, thus enabling the next round of the EF-G cycle to take place.

The situation with EF-Tu is more complicated because the binary complex EF-Tu-GTP cannot bind to the ribosome. This process can be accomplished only by the ternary complex aminoacyl-tRNA-EF-Tu-GTP, marking a second main function of EF-Tu as well as promotion of the transition from the post- to the pretranslocation state — namely to be a carrier which brings aminoacyl-tRNAs to the ribosomal A site.

The crystal structure of the ternary complex<sup>5</sup> shows that EF-Tu only contacts the aminoacyl-tRNA far away from the anticodon stem-loop structure at a few nucleotides at the T stem and the acceptor arm including the aminoacyl residue. Intriguingly, this means that the decoding process (decoding step in the figure) possibly does not involve EF-Tu but rather is confined to the interaction between the anticodon and the codon at the ribosomal decoding centre  $\delta$ . In this view the high binding energy of both EF-Tu and tRNA outside the anticodon region is exploited to drive the transition to the pre-translocational state only after decoding has finished successfully (see figure — the step at which the A position is occupied).

Comparison of the structure of the ternary complex with the structures of EF-G in the GDP-bound<sup>6</sup> or in the nucleotide-free state<sup>7</sup> revealed an exciting similarity between the tRNA part of the ternary complex and the domains 3 to 5 of EF-G. This molecular mimicry of a portion of EF-G might be explained in the following way. After EF-G has established the post-translocational state, its



The ribosomal elongation cycle with the two GTPase cycles of EF-Tu and EF-G. The GTPase cycles are described in the text. The ribosome translates genetic information into protein structure by transforming a sequence of codons on the messenger RNA into the sequence of the corresponding amino acids. The ribosomes use transfer RNAs for this translational process. The tRNAs expose an anticodon at one end and carry the corresponding amino acid at the other end. In the course of the ribosomal elongation cycle, two tRNAs are always on the ribosome. A, P and E are the places where a tRNA can show up. The A region (for aminoacyl-tRNA) contains the decoding centre  $\delta$  where the match of codon and anticodon of the newly incoming aminoacyl-tRNA is checked. The P region houses the peptidyl-tRNA before the peptidyl residue is transferred to the aminoacyl-tRNA at the A region. At the E region only deacylated tRNA is found; from here the deacylated tRNA is released from the ribosome. The ribosomal elongation cycle is presented in the frame of the  $\alpha$ - $\epsilon$  model according to which both tRNAs are tightly bound to a movable domain which contains two binding sites  $\alpha$  and  $\epsilon$ . The movable domain displays the tRNAs at the A and P regions before translocation and at the P and E regions after translocation. The decoding centre  $\delta$  is not movable but rather fixed at the A region. It superposes the  $\alpha$  site in the pre-translocation state (Pre) but stands alone in the post-translocational state (Post).

tRNA-like structure might enter the decoding centre  $\delta$  at the A position, thus blocking a reverse translocation as long as the EF-G-GTP complex is on the ribosome keeping the activation-energy barrier low. As soon as EF-G-GDP has left the ribosome the activation-energy barrier is raised again and the ribosome is trapped in the post-translocational state until the advent of the next cognate ternary complex.

When the occupation of the A position is completed, the freshly established pre-translocational state activates the intrinsic GTPase centre of EF-Tu and EF-Tu-GDP leaves the ribosome; here again the high energy barrier traps the ribosome in a defined state, this time the pre-translocational state. In striking contrast to EF-G-GDP, the affinity of EF-Tu for GDP or GTP is so high that it cannot be regulated by their cytoplasmic concentrations. EF-Tu therefore requires an additional factor which allows the exchange of the guanine nucleotide. This factor is the protein EF-Ts.

The structure of the EF-Tu-EF-Ts complex described by Kawashima *et al.*<sup>8</sup> surprisingly reveals not a heterodimer, but a tetramer, in which both EF-Ts molecules

make tight contacts and the EF-Tu molecules hardly touch each other. The complex can therefore best be described with the formula EF-Tu·(EF-Ts)<sub>2</sub>·EF-Tu. In the complex, EF-Tu has a conformation which is related to that of EF-Tu·GDP.

We can now understand why the complex results in a release of the tightly bound GDP. An EF-Ts sequence threonine-aspartic-acid-phenylalanine-valine, conserved in all species, plays a central role in this release: the aspartic acid and phenylalanine at positions 80 and 81 press into the GDP-binding site and displace three amino acids of EF-Tu which are stabilizing two water molecules. Both water molecules are coordinated to a Mg<sup>2+</sup> ion which is essential for the affinity to GDP. The disruption of the Mg<sup>2+</sup>-binding site leads to a loss of GDP, which now can be replaced by GTP. The replacements are reversible, but for an efficient GTPase cycle GTP must replace GDP rather than vice versa. The step disturbing this system of equilibria is the binding of the aminoacyl-tRNA which sequesters EF-Tu·GTP<sup>10</sup> and drives the EF-Tu cycle forward (see figure).

We now broadly understand the struc-

tural changes of EF-Tu at the molecular level, and the same will be true for the EF-G cycle if the structure of EF-G·GTP can be solved. Only the ribosome keeps its secrets, leaving the questions of how the elongation factors promote the ribosomal transitions between the pre- and the post-translocational states, and of the mechanism by which the ribosome triggers the GTP cleavage on the factors. □

Knud H. Nierhaus is in the Max-Planck-Institut für Molekulare Genetik, Ihnestr. 73, D-14195 Berlin, Germany.

1. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117–127 (1991).
2. Kjeldgaard, M. & Nyborg, J. *J. molec. Biol.* **223**, 721–742 (1992).
3. Berchtold, H. *et al. Nature* **365**, 126–132 (1993).
4. Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. *Structure* **1**, 35–50 (1993).
5. Nissen, P. *et al. Science* **270**, 1464–1472 (1995).
6. Czerwowski, J., Wang, T. J., Steitz, T. A. & Moore, P. B. *EMBO J.* **13**, 3661–3668 (1994).
7. Evarsson, A. *et al. EMBO J.* **13**, 3669–3677 (1994).
8. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. *Nature* **379**, 511–518 (1996).
9. Schilling-Bartetzko, S., Bartetzko, A. & Nierhaus, K. H. *J. biol. Chem.* **267**, 4703–4712 (1992).
10. Romero, G., Chau, V. & Biltonen, R. L. *J. biol. Chem.* **260**, 6167–6174 (1985).

## BIOGEOGRAPHY

### Vent fauna and plate tectonics

As any experienced traveller will know, the most direct route from A to B is not always the quickest. On page 531, V. Tunnicliffe and C. M. R. Fowler show that the ancestors of the spider crab and tube worms pictured here kept in touch via the highways of the global mid-ocean ridge system, rather than by crossing the open country of the abyssal plains. The results explain some otherwise puzzling features of the biogeography of deep-sea hydrothermal vent communities, and offer the tantalizing possibility that their further study might shed light on past configurations of the Earth's tectonic plates.



The discovery of the first hydrothermal vent community in 1977 (on the Galapagos Rift, off the coast of Ecuador) came as a complete surprise — first, for the sheer abundance of organisms thriving in the midst of the 'desert' of the deep sea, but then too for the nature of the newly discovered creatures. Of the 375 species that have now been described from vents worldwide, over 90 per cent are restricted to the vent habitat. This strong 'endemicity' reflects the rather special adaptations required to live in such strongly reducing, metal-rich surroundings.

From a biogeographical point of view, the most salient characteristic of the

vent habitat is its extreme discontinuity, in both space and time. Active vent fields may be separated by hundreds of kilometres, with the intervening stretches of mid-ocean ridge as unwelcoming to life as the rest of the deep sea. And the good times don't last: the hydrothermal activity that provides the chemical energy source for a vent community may persist for only a few years before petering out, leaving the vent organisms stranded. As most of the vent inhabitants are sessile or slowly moving, they depend on their larvae to find and colonize new vents. This reliance on larval dispersal to propagate the species is reflected in the provinciality of vent

fauna: whereas two sites about 800 km apart on the East Pacific Rise have 54 species in common, only five species are shared between the western and eastern Pacific, and only one between the Pacific and Atlantic oceans.

Tunnicliffe and Fowler propose that species migration should occur most easily along mid-ocean ridges, where there is at least a chance of finding a congenial rest stop on the journey. (Others have proposed that rotting whale carcasses might provide suitable stepping stones across the open ocean, but they are at best imperfect substitutes for vents.) Data from eight vents in the Pacific and Atlantic support this proposal: the distance between two vents measured along the mid-ocean ridge system is a better predictor of their species similarity than is the direct, cross-country distance. Moreover, the biogeographical relationships seem to bear the imprint, not just of today's ridge connections, but of those that existed tens of millions of years ago. It may not be too fanciful to suggest that, when vent faunas from the rest of the world's oceans have been found and classified, their phylogeny and distribution may provide information about past plate geometries that the processes of plate tectonics have long since erased. Laura Garwin



# Fertilization without sperm

R. John Aitken and D. Stewart Irvine

INFERTILITY affects about one in seven couples in the United Kingdom, and in about half of these cases it is defects in the male partner that cause the failure to conceive. Within the past year, however, new assisted-conception techniques mean that immature spermatozoa taken from testes of men with the severest forms of reproductive dysfunction can achieve both fertilization and normal development of the embryo. Even precursor germ cells (spermatids and spermatocytes) that have not yet differentiated into spermatozoa have been used to accomplish the same end. Behind these astonishing feats lies work from several laboratories, including those of Yanagimachi and co-workers<sup>1,2</sup>, Tesarik *et al.*<sup>3</sup> and Fishel *et al.*<sup>4</sup>.

These advances stem from micro-manipulation techniques for injecting individual male gametes directly into the egg cytoplasm, thereby bypassing the very complex series of cell-cell interactions that mediate normal fertilization. The practitioners in this field have been so ingenious and successful, that their research has raised ethical questions about the lengths to which artificial means should be used for human procreation, about the rights of couples to have children and about the rights of those children to be normal.

What is the basic biology behind these embryological achievements? Under normal circumstances, spermatogenesis results in the generation of  $1n$  haploid spermatozoa, where  $n$  represents the normal DNA content of a mature cell and haploid means that each contains a single set of 23 chromosomes. In contrast, arrest in metaphase of the second meiotic division gives the ovulated egg a  $2n$  haploid constitution (called a metaphase-2 oocyte) until the oocyte is activated at fertilization (Fig. 1). The latter involves a complicated series of interactive events that can fail at any stage. For spermatozoa lacking the vigorous motility necessary to penetrate the zona pellucida, procedures such as partial zona dissection, in which penetration is achieved through a breach in the zona pellucida, or subzonal insemination<sup>5</sup>, in which spermatozoa are injected directly into the perivitelline space, were developed to facilitate fertilization.

Although such techniques enjoyed moderate success, they did not help men whose spermatozoa were incapable of initiating fusion with the oocyte. For such patients, the only solution was to inject a single spermatozoon directly into the ooplasm. This technique (intracytoplasmic sperm injection or ICSI)<sup>6</sup> not only worked, but worked well<sup>7</sup>. Patients producing vanishingly small numbers of spermatozoa

could be helped, and the pregnancy rates obtained were as good, if not better, than most centres achieve with *in vitro* fertilization when semen quality is normal<sup>7</sup>.

The requirement that spermatozoa should be present in the ejaculate was overridden when it was discovered that immature spermatozoa from the epididymis could initiate normal pregnancies following their injection into the oocyte<sup>8,9</sup>. With the aid of such techniques, men with blockage of the excurrent duct system due to congenital absence of the vas deferens, or vasectomy, can be treated successfully. Later studies showed that it was even unnecessary for spermatozoa to reach the excurrent ducts, because testicular spermatozoa could induce normal embryonic development following microinjection into the oocyte<sup>10,11</sup>. And now the need for spermatogenesis to have proceeded as far as the differentiation of a spermatozoon is obviated by the latest experiments involving the use of precursor germ cells in assisted fertilization procedures<sup>1-4</sup>.

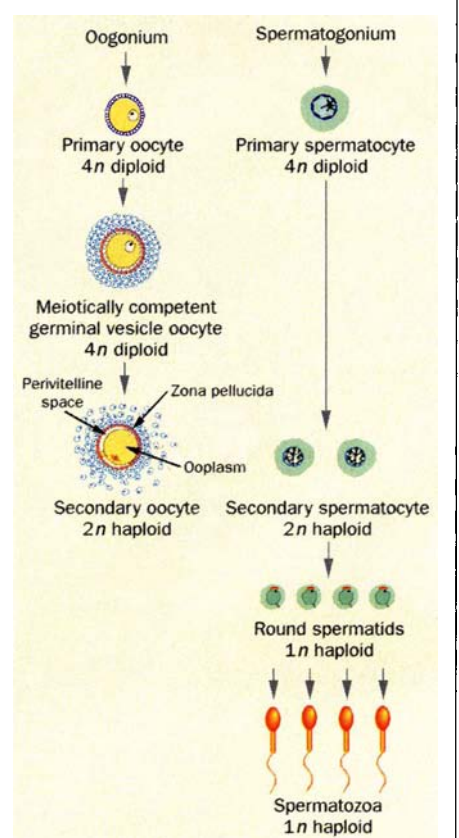
Ogura *et al.*<sup>1</sup> showed that, in mice,  $1n$  haploid round spermatids can be fused with mature, metaphase-2 oocytes and initiate normal embryonic development. This

remarkable experiment demonstrated that all of the postmeiotic modifications that result in the differentiation of a mature spermatozoon serve only to deliver a competent, haploid, male genome to the oocyte. In the mouse, the only function performed by the mature spermatozoon that cannot be replicated by a spermatid is induction of the calcium wave that activates the oocyte. The authors replaced this particular sperm function by giving the ovum a series of electric shocks.

These findings also have implications for the imprinting of the male genome. This process is thought to involve cytosine methylation and follows a complementary, gender-specific pattern such that normal embryonic development can only occur if fertilization involves a male and a female gamete. Generation of a diploid zygote through the fusion of two haploid oocytes or the penetration of an enucleated egg with two haploid spermatozoa does not give rise to a viable embryo. The fact that spermatid injection can support normal embryonic development must mean that the male genome is imprinted (or has gained the potential to imprint) by this stage of spermatogenesis.

Clinically, these results also suggest that treatments involving spermatids might be of value for azoospermic patients in whom the differentiation of spermatozoa from spermatid precursors

FIG. 1 Meiosis in the male and female gamete. During oogenesis, meiosis is initiated in the fetus but becomes arrested at the diplotene stage of first meiotic prophase. At this stage the primary oocyte exists as a primordial follicle with a  $4n$  diploid chromosome complement contained within a germinal vesicle. At birth, all germ cells are arrested at this stage. No new oocytes are generated in later life, so that the store of primordial follicles present in the neonatal ovary represents the entire lifetime's supply of ova. In the human, primordial follicles may remain in this arrested state for 30–40 years before being recruited into the growing follicle population. During folliculogenesis, the oocyte grows but remains blocked at the first meiotic prophase. It is only at ovulation that the arrest is temporarily lifted and the oocyte progresses to metaphase 2, discarding half of its chromosomes in the first polar body to render the secondary oocyte  $2n$  haploid. Meiosis in the female gamete is finally completed at fertilization, when an intracellular calcium wave generated in response to sperm penetration precipitates the extrusion of the second polar body and formation of the  $1n$  haploid female pronucleus. Gametogenesis in the male is an uninterrupted process initiated at puberty and continuing throughout reproductive life. Meiosis produces  $1n$  haploid spermatids that then undergo a remarkable transformation into fully differentiated spermatozoa. Subsequently, the latter have to undergo a period of maturation in the epididymis, followed by a phase of capacitation in the female reproductive tract, before they acquire the ability to fertilize the ovum.



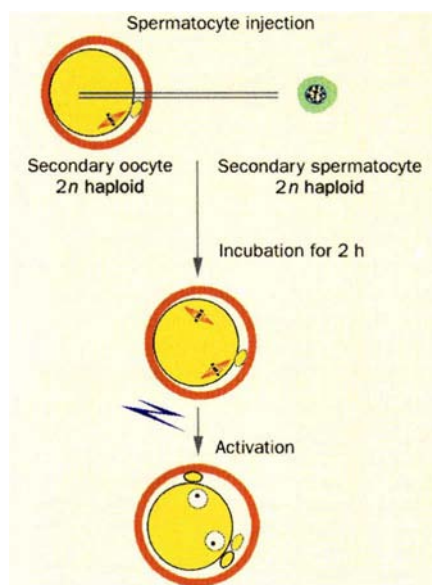


FIG. 2 Spermatocyte injection. The technique described by Kimura and Yanagimachi<sup>2</sup> involved the injection of a 2n haploid secondary spermatocyte nucleus into a mature 2n haploid oocyte in metaphase 2. The spermatocyte chromosomes responded to signals in the ooplasm by entering metaphase during a 2-hour incubation period. When the oocyte was subsequently electroactivated, both the male and female chromosomes resumed meiosis, extruded a polar body, formed a pronucleus and initiated normal embryonic development.

(spermiogenesis) is defective. The most recent results<sup>3,4</sup> indicate that this objective has now been achieved, with at least one normal infant being born to an azoospermic father following the intracytoplasmic injection of a spermatid. That this procedure is effective in humans is particularly interesting because in man, but not the mouse, the sperm centrosome plays a key role in microtubule organization during cell division. It is therefore clear that the paternal centrosome acquires the capacity to orchestrate cell division in the zygote, before the differentiation of a spermatozoon.

Genetically, the fertilization of an egg with a spermatid rather than a spermatozoon is understandable because both cell types are in the 1n haploid state (Fig. 1). Much more remarkable is the description<sup>3</sup> of the successful development of viable progeny following the fertilization of metaphase-2 oocytes with secondary spermatocyte nuclei which, although they are haploid, contain twice the haploid amount of DNA (2n haploid). Interestingly, the secondary spermatocyte is in exactly the same genetic state as the ovulated egg (Fig. 1) and could be induced to complete its meiosis in synchrony with the latter<sup>2</sup>. Thus, following injection, the secondary spermatocyte nucleus was found to respond to the presence of factors in the ooplasm (maturation-promoting factor and cytostatic factor) by undergoing chro-

mosome condensation and forming a metaphase plate (Fig. 2). When the oocyte was electroactivated, both the male and the female metaphase chromosomes synchronously entered anaphase, extruded a polar body and formed a pronucleus. The resulting zygotes were able to develop to term on transplantation to foster mothers.

Inevitably, these techniques will be used to examine the developmental potential of cells derived from earlier stages of spermatogenesis, while the genome is still in a diploid state (that is, with both sets of chromosomes, one from each parent). Such studies should not only extend our knowledge of genomic imprinting but also indicate the competence of the male genome to undergo meiotic arrest in response to signals provided by the oocyte. They also open up the possibility of assisted conception therapy for patients whose spermatogenesis cannot progress beyond the secondary spermatocyte stage. But it should be recognized that the paternal centrosome inheritance characteristic of human fertilization will complicate matters, because it is not only a viable nucleus that must be transferred to the oocyte but a functional centrosome as well.

The new techniques are not without critics, on scientific as well as ethical grounds. For example, although the health of children conceived after ICSI is broadly reassuring<sup>7</sup>, some legitimate concerns have been raised. The disruption of spermatogenesis in some infertile patients with non-obstructive azoospermia will have a genetic origin involving, for example,

microdeletions in the Y chromosome<sup>12,13</sup> or sex chromosome abnormalities such as XXY. Even in cases of obstructive azoospermia genetic factors may be involved, as evidenced by the association between congenital absence of the vas deferens and mutations in the cystic fibrosis gene<sup>14</sup>. The indiscriminate use of ICSI may therefore lead to the transmission of genes that may either perpetuate the infertility or give rise to complications with a related genetic background. It is only when the causes of male infertility are better known that effective screening for genetic factors predisposing to the condition will be feasible, and that we will have a rational basis for selecting and counselling patients for this novel form of therapy. □

R. John Aitken and D. Stewart Irvine are in the MRC Reproductive Biology Unit, Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh EH3 9EW, UK.

1. Ogura, A., Matsuda, J. & Yanagimachi, R. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7460–7462 (1994).
2. Kimura, Y. & Yanagimachi, R. *Biol. Reprod.* **53**, 855–862 (1995).
3. Tesarik, J., Mendoza, C. & Testart, J. *New Engl. J. Med.* **333**, 525 (1995).
4. Fishel, S. et al. *Lancet* **345**, 1641–1642 (1995).
5. Fishel, S., Timson, J., Lisi, F. & Rinaldi, L. *Fertil. Steril.* **57**, 840–849 (1992).
6. Palermo, G. et al. *Lancet* **340**, 17–18 (1992).
7. Van Steirteghem, A. et al. *Hum. Reprod.* **10**, 2527–2528 (1995).
8. Tournaye, H. et al. *Fertil. Steril.* **61**, 1045–1051 (1994).
9. Devroey, P. et al. *Hum. Reprod.* **10**, 903–906 (1995).
10. Silber, S. et al. *Hum. Reprod.* **10**, 148–152 (1995).
11. Devroey, P. et al. *Hum. Reprod.* **10**, 1457–1460 (1995).
12. Mam, K. et al. *Hum. molec. Genet.* **1**, 29–33 (1992).
13. Kobayashi, K. et al. *Hum. molec. Genet.* **3**, 1965–1967 (1994).
14. Patrizio, P., Asch, R. H., Handelin, B. & Silber, S. *Hum. Reprod.* **8**, 215–220 (1993).

## PLANETARY SCIENCE

# The subtle taste of Jupiter

David J. Stevenson

PLANETS, like wines, mature with time. On 7 December last year, a probe from the Galileo spacecraft tasted Jupiter, revealing that at least in one place and epoch the planet is flavoured less richly than many expected. Preliminary results, discussed at a news conference on 22 January, reveal lower amounts of oxygen (as water), sulphur, helium and neon than presupposed, all measured relative to the overwhelmingly abundant hydrogen. Tiny but possibly high levels of xenon and krypton complicate the picture. But in all these subtle additives lies a wealth of information about how Jupiter and like planets form and evolve.

Galileo is set to study Jupiter's system over the next two years, but its initial goal was to release a probe into the atmosphere. During a 58-minute-long descent, data were sent from the probe to the orbiter, and later back to Earth. They cover a depth of about 159 km and begin

at roughly the ammonia cirrus cloud level one sees when looking at Jupiter through a telescope. The pressure and temperature increased from about 0.4 bar and 129 K to about 22 bar and 425 K, where transmission ceased.

First some background. Jupiter is mostly hydrogen<sup>1</sup> and is very hot internally (it emits twice as much energy as it receives from the Sun<sup>2</sup>). It has no surface in the normal sense, because the temperature–pressure profile corresponding to the mostly convective heat outflow always ensures that the planetary material is far too hot to liquefy or solidify, at least until the very deep interior. The mostly convective state also means that the atmosphere and interior can probably exchange material.

Ideas about how the planet formed have several features in common<sup>3</sup>. In the disk of matter orbiting the proto-Sun, some 4.5 billion years ago, rock and water



ice coagulated to form planetesimals. In the Jupiter-forming zone, enough of this material (perhaps several Earth masses) somehow concentrated in one body. Gravitational influence then led to a steady inflow of gas, building up the planet that we now see. The gas envelope consisted mainly of hydrogen and helium, with an overall composition that was enriched in heavier elements. This story is compatible with models of Jupiter as it is now<sup>4</sup>, although one might expect the atmosphere to be depleted in water (because the water is in the core) or any other condensate. However, it may have been that planetesimals still entering Jupiter during and after the gas inflow broke up, their constituents mixing with the gas<sup>3,5</sup>. Convective mixing from the interior might also bring heavier elements up from the core. Support for enrichment comes from observations of carbon (as methane) in the atmosphere<sup>5</sup>.

Noble gases present a different set of issues. Helium and neon do not condense or stick to solid particles at the temperatures concerned, so it is reasonable to suppose that they were initially in cosmic proportions in the atmosphere. However, helium is believed to be insoluble in hydrogen at very high pressures, even at very high temperatures, because the hydrogen metallizes, and the interaction between light noble gases and conduction electrons is highly repulsive<sup>6</sup>. As Jupiter cooled over time, runs the argument, helium would 'rain out' in the deep interior, and because of that and convective mixing the atmosphere would become depleted in the element. Neon is far less abundant and thus unlikely to form its own raindrops, but would partition into the helium raindrops and perhaps undergo even larger depletion than helium<sup>7</sup>. Heavy noble gases, on the other hand, stick to solid particles or form inclusion compounds with water ice, and are moreover likely to be far more soluble in a metallic medium.

So to the Galileo results, which reveal that the oxygen to hydrogen ratio at the deepest regions sampled (below the expected depth of water cloud formation) is about the same as in the Sun. Sulphur is somewhat lower than the solar mixing ratio. Neon is markedly less than solar (by perhaps two orders of magnitude), and helium is 13.7% by mass (about half the primordial solar value). But krypton and especially xenon may substantially exceed solar abundance. The dryness of the atmosphere is compatible with the low cloudiness measured by the probe, the higher temperatures than predicted (actually suggesting even lower water abundances than a solar mixing ratio), and the net flux radiometric measurements.

So what can we deduce from these data? Lower than expected water might mean that the planetesimals hitting

Jupiter early in its history had a carbon to oxygen ratio in excess of solar. Or perhaps the water depletion is just a meteorological effect peculiar to the place that the probe went in. Or there might even be some continuing process within the planet that depletes oxygen.

The unexpectedly low levels of helium must be reconciled with the energy output of the planet, because helium rainout releases huge amounts of energy<sup>8,9</sup>. Perhaps the depletion begins either higher up (for example at one megabar pressure or less) or deeper down (many megabar) than was supposed. In the first case, a smaller mass-fraction of the planet undergoes depletion (maybe only 10%); in the second, the helium falls less distance. Perhaps the deepest part of the planet is stably stratified and not releasing heat by convection. In any event, neon and perhaps some argon will go along for the ride in helium raindrops. The lower abundance of helium might also lend support to the existence of a deep radiative zone<sup>10</sup>, because the outer envelope is otherwise not dense enough to be compatible with the gravity field. The high xenon is perhaps the hardest to explain, but it could result from delivery by planetesimals and some internal partitioning within the planet. Sharpening the speculations will require more precise data than released thus far, as well as the incorporation of the latest ideas of how hydrogen may metallize<sup>11</sup>.

The probe also recorded high east-west winds, even at the greatest depths sampled, less lightning than many expected, and an inner radiation belt inside Jupiter's dust ring. In all, the observations will keep scientists occupied until well after the excitement picks up anew when Galileo starts to encounter the planet's satellites, beginning with Ganymede next June. □

David J. Stevenson is in the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA.

1. Wildt, R. in *The Solar System*, Vol. 3, *Planets and Satellites* (eds Kuiper, G. P. & Middlehurst, B. M.) 159–212 (Univ. Chicago Press, 1961).
2. Low, F. J. *Lowell Observ. Bull.* **129**, 184–187 (1965).
3. Pollack, J. B. & Bodenheimer, P. in *Origin and Evolution of Planetary and Satellite Atmospheres* (eds Atreya, S. K., Pollack, J. B. & Matthews, M. S.) 564–604 (Univ. Arizona Press, Tucson, 1989).
4. Zharkov, V. N. & Gudkova, T. V. *Ann. Geophysicae* **9**, 357–366 (1991).
5. Gautier, D. & Owen, T. in *Origin and Evolution of Planetary and Satellite Atmospheres* (eds Atreya, S. K., Pollack, J. B. & Matthews, M. S.) 487–512 (Univ. Arizona Press, Tucson, 1989).
6. Stevenson, D. J. *J. Phys. F. (Metal Physics)* **9**, 791–801 (1979).
7. Roulston, M. S. & Stevenson, D. J. *EOS (Abstr. Fall Meeting)* 343 (1995).
8. Smoluchowski, R. *Nature* **215**, 691–695 (1967).
9. Salpeter, E. E. *Astrophys. J. Lett.* **181**, L183–L186 (1973).
10. Guillot, T., Gautier, D., Chabrier, G. & Mosser, B. *Icarus* **112**, 337–353 (1994).
11. Weir, S. T., Mitchell, A. C. & Nellis, W. J. *Phys. Rev. Lett.* (in the press).

## Small and slippery

In large chunks, graphite has a planar layered structure. For small chunks, this laminar form is no longer optimal. A spherical 'onion' form is stabler; the absence of dangling edge-bonds more than compensates for the strain energy needed to bend the carbon sheets. Soot can contain spherical sub-micrometre 'buckyonions' of nested fullerene shells, and spherical graphite inclusions as large as 0.1 mm across occur in cast iron.

Daedalus now wants to create an alloy with sub-micrometre spherical inclusions which rotate freely in their metal matrix. Where they protrude from the surface they will act as ball-bearings, almost abolishing skin friction. Buckyonions seem ideal for this job; the field is advancing so fast that they should soon be available in quantity. Daedalus plans to include them, not in iron, but in a metal lattice carefully designed to be slightly mismatched to their peripheral atomic spacing. Each buckyunion will then have no preferred orientation in its metal matrix. In any orientation some regions on its surface will match the matrix atom-to-atom, while in other regions its atoms will coincide with matrix interatomic spaces. This is the Merkle principle for a nanotechnological bearing. By mismatching the atomic spacing of shaft and bearing, lattice 'cogging' is overcome, and the shaft spins freely.

The resulting alloy will be far more slippery than ice. Its obvious use will be in machine tools, drills, and cutters generally. Almost all the power needed for cutting and machining, and all the forces to be withstood, arise from the rubbing of the sheared metal against the cutting tool. But Daedalus's 'Buckalloy' tools will offer no friction to the sliding, deforming metal. It will ride smoothly on innumerable free-spinning buckyonions, each protruding a little from the surface of the tool. Strong and smooth as carbon fibre, they will carry the heaviest load uncomplainingly.

This cunning surface will resist wear, too. Even if it slowly abrades away, new buckyonions will be exposed as fast as old ones are undercut and fall out. By the same token, Buckalloy will be hard to shape or grind. It will have to be cast and chemically etched into shape.

Once it has won its spurs as a pricey tool steel, Buckalloy will slide down-market into wider low-friction duties. Wonderfully smooth heavy-duty bearings, slides, gears and pivots will be possible, each with a Buckalloy surface rolling against a hard smooth steel one, and none of them needing any oil. Freed from the stigma of oily hands, engineers may even attain the social status of stockbrokers.

David Jones

# Recognizing facial emotion

**SIR**—The amygdala has been implicated in emotion and emotional memory in humans and experimental animals<sup>1,2</sup>. Some of us<sup>3</sup> have reported impaired recognition of emotion in facial expressions (especially fear) in a 30-year-old woman (S.M.) with a nearly complete bilateral lesion of the amygdala due to congenital Urbach-Wiethe disease. This suggested that the amygdala may be essential not only for emotional learning, but also for recognition of emotion in facial expressions. Here we report data from two men who survived herpes simplex encephalitis (E.P.<sup>4</sup>

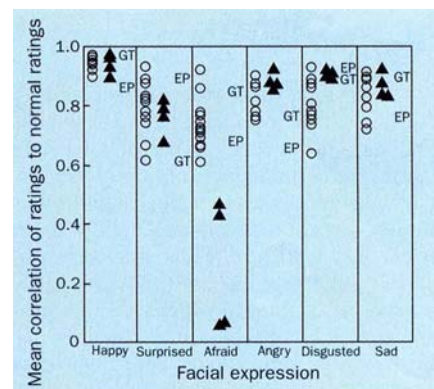
and G.T., aged 73 and 59 years, respectively), with complete bilateral lesions of the amygdala and additional temporal lobe structures (Fig. 1). Despite this, they appear unimpaired in recognizing facial expressions of emotion, including fear. E.P. and G.T. are profoundly amnesic but have IQ scores in the normal range.

E.P. and G.T. were tested twice, once using the exact materials and procedure reported by Adolphs *et al.*<sup>3</sup> (Fig. 2), and again using a slightly modified version of the same experiment. E.P. and G.T. rated normally the same facial expressions that S.M. rated abnormally. For the modified test, in which the emotional adjectives varied randomly from trial to trial, the two patients performed similarly to four age- and education-matched controls in every respect (correlations between groups >0.70 for each of the six emotions). For faces consistent with the emotion being rated, the patients obtained an average rating of 4.6 (4.8 for fear) on a 0–5 scale; control rating was 4.3 (4.3 for fear).

There are two possibilities consistent with these data. One possibility is that bilateral damage to the amygdala together with damage outside the amygdala is required to impair the recognition of facial emotion. Intracranial lesions resulting from Urbach-Wiethe disease can extend beyond the amygdala to include other structures<sup>5,6</sup>. In addition to her amygdala lesion, S.M. has minimal damage to the entorhinal cortex<sup>2</sup> on the left side. Moreover, postmortem histopathological analysis may reveal damage not detected by current neuroimaging techniques. Damage to the basal ganglia is reported to impair the recognition of facial emotion<sup>7</sup>, although it is unclear that such impairment is similar to that of S.M.

A more likely possibility is that amygdala lesions impair the recognition of emotion in facial expressions only if these lesions occur early in development, rather than in adulthood. The Urbach-Wiethe disease that led to S.M.'s amygdaloid lesions is congenital, whereas E.P. and G.T. sustained their lesions after the age of 50 years. D.R., a 51-year-old woman who is also deficient in recognizing facial emotion<sup>8</sup>, sustained partial bilateral lesions to the amygdala and to the right basal ganglia in adulthood. However, D.R. has a history of epilepsy and a broader deficit than S.M. in processing emotions. Perhaps a combination of factors, including a congenital lesion, low full-scale IQ (S.M. = 86, D.R. = 87) and/or additional brain damage, could determine how readily other strategies are available for recognizing facial emotion.

The present report suggests that in the



**FIG. 2** Correlations of ratings of facial expressions on emotional adjectives with ratings given by previously studied<sup>3</sup> normal controls ( $n=7$ ). Mean Pearson correlations are shown for the ratings given by brain-damaged controls (circles;  $n=12$ ), S.M.'s ratings (triangles show results for 4 experiments), E.P.'s ratings (EP; 1 expt), and G.T.'s ratings (GT; 1 expt). Data for the brain-damaged controls and S.M. are from Adolphs *et al.*<sup>3</sup>. Data for E.P. and G.T. are from an experiment that replicated the procedures used therein. Recognition of facial expression of human emotion was tested by presenting on a screen and one at a time 6 different expressions of facial emotion (each displayed by 6 different faces). Subjects rated each face with respect to 9 adjectives (the 6 emotions shown in the figure, as well as awake, sleepy or interested).

absence of damage early in life, recognition of facial emotion in adults does not have an absolute dependence on the amygdala. These data and conclusions are fully compatible with an important role for the amygdala in the experience and expression of emotion. Tasks that assess the recognition of fear in facial expressions may not require the actual experience of fear.

**Stephan B. Hamann, Lisa Stefanacci**

Department of Psychiatry,  
University of California,  
San Diego, California 92093, USA

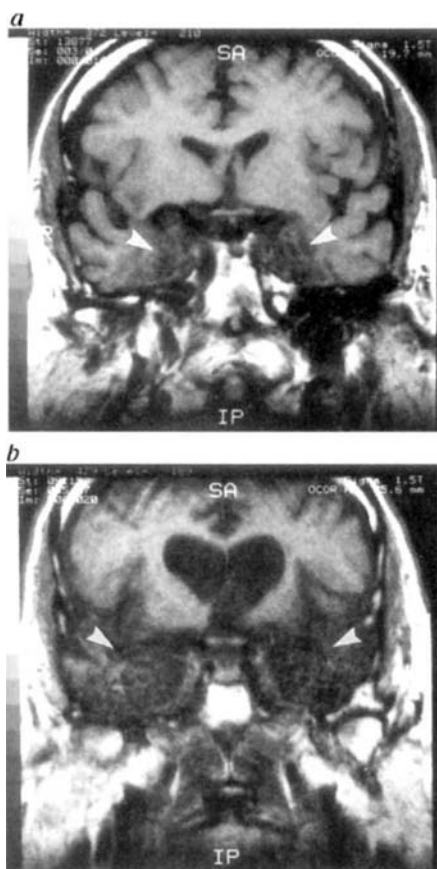
**Larry R. Squire**

Departments of Psychiatry  
and Neurosciences,  
University of California,  
San Diego, California 92093, USA and  
VA Medical Center,  
San Diego, California 92161, USA

**Ralph Adolphs, Daniel Tranel**

Hanna Damasio, Antonio Damasio  
Department of Neurology,  
University of Iowa, College of Medicine,  
Iowa City, Iowa 52242, USA

1. LeDoux, J. *Behav. Brain Res.* **58**, 69–79 (1993).
2. Bechara, A. *et al. Science* **269**, 1115–1118 (1995).
3. Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. *Nature* **372**, 669–672 (1994).
4. Squire, L. & Knowlton, B. *Proc. natn. Acad. Sci. U.S.A.* **92**, 12470–12474 (1995).
5. Friedman, D., Mathews, R. & Swanepoel, P. *S. Afr. med. J.* **65**, 734–735 (1984).
6. Newton, F., Rosenberg, R., Lampert, P. & O'Brien, J. *Neurology* **21**, 1205–1213 (1971).
7. Jacobs, D., Shuren, J., Bowers, D. & Heilman, K. *Neurology* **45**, 1696–1702 (1995).
8. Young, A., Hellawell, D., van de Wal, C. & Johnson, M. *Neuropsychologia* (in the press).



**FIG. 1** *a*, Coronal T1-weighted magnetic resonance image of patient E.P. showing extensive bilateral damage to the amygdaloid complex (arrows). The damage extends rostrally to the temporal pole and caudally to include the hippocampal formation. There is also severe atrophy of the entorhinal, perirhinal and parahippocampal cortices. *b*, Coronal T1-weighted magnetic resonance image of patient G.T. at a similar rostrocaudal level, showing severe bilateral temporal lobe damage (arrows). The damage extends through the anterior 7.0 cm of G.T.'s left temporal lobe and through the anterior 5.0 cm of his right temporal lobe. The lesion bilaterally includes the amygdaloid complex, hippocampus, entorhinal, perirhinal and parahippocampal cortices as well as the inferior, middle and superior temporal gyri. There is also bilateral damage in the insular cortex, medial and orbital frontal cortex, and cingulate gyrus.

## Kin recognition in honeybees

**SIR**—In honeybee colonies, polyandry leads to the presence of 7–20 subfamilies<sup>1</sup>. Workers can discriminate members of the same subfamily (super-sisters) from workers of other subfamilies (half-sisters)<sup>2,3</sup>, and this may allow them to act nepotistically. How they discriminate is unknown, but the process is likely to involve subfamily-specific chemical labels which bees learn<sup>2</sup>. Page *et al.*<sup>4</sup> showed that laboratory-reared workers from artificial colonies with just two subfamilies had cuticle hydrocarbon profiles that were more similar between super-sisters than between half-sisters. This may not reflect the natural situation, however, where nest-mates can exchange hydrocarbons either by direct contact or through comb wax.

We have examined the cuticle hydrocarbon composition of honeybee workers from a colony headed by a naturally inseminated queen. We assigned each bee to one of 16 subfamilies using two highly variable microsatellite loci (*A76*, *A107*; ref. 1), genetic markers unrelated to cuticle hydrocarbons. To determine the relative importance of genetic and environmental factors to cuticle hydrocarbon profiles, adult bees were matured under three conditions: isolated, in groups of 10, and in their parental hive. After 5 days, workers of each set (117, 77 and 117 bees, respectively) were individually analysed (see figure).

Hydrocarbons were extracted for 5 minutes in 1 ml of pentane, analysed by gas chromatography (Girdel 300) on a 30-m nonpolar capillary column and constituents identified by mass spectrometry

(Nermag R 10-10-C GC-MS). Twenty-six compounds were identified, belonging to four major classes of long-chain hydrocarbons: alkenes, alkadienes, methyl-branched alkanes, but mostly *n*-alkanes. All *n*-alkanes in the  $C_{21}$ – $C_{33}$  series were present. Compounds with an even number of carbons predominated. For the statistical analyses, 12 compounds were excluded because their concentrations were too low to be measured reliably. The 14 remaining compounds were  $C_{23}$ ,  $C_{23:1}$ ,  $C_{25}$ ,  $C_{27}$ ,  $C_{27:1}$ ,  $MeC_{27}$ ,  $C_{29}$ ,  $C_{29:1}$ ,  $MeC_{29}$ ,  $C_{31}$ ,  $C_{31:1}$ ,  $MeC_{31}$ ,  $C_{33:1}$ ,  $C_{33:2}$ .

A generalized linear model<sup>5</sup> (Splu 3.0) was applied to the data matrix. Each element of this matrix is the mean percentage of a given hydrocarbon for a family in a given rearing condition. Each matrix line thus represents a mean hydrocarbon profile. Our analysis shows that hydrocarbon profiles differ significantly between subfamilies. This demonstrates that these profiles are conserved even within the hive, and that cuticle hydrocarbons possess the necessary prerequisites of sufficient variability and genetic determinism for use as labels for subfamily recognition.

A potential consequence of subfamily recognition is that workers could improve the reproductive success of their own subfamily in several circumstances, such as rearing full-sister queen larvae<sup>6</sup> or preferentially feeding their full-sister laying workers. Such nepotistic behaviour, claimed by some<sup>6,7</sup> but denied by others<sup>8,9</sup>, implies subfamily recognition. Breed *et al.*<sup>10</sup> argued that few areas in sociobiology have received as much experimental attention, yet yielded so little in the way of supportable conclusions, as the question of subfamily nepotism in honeybees. Our demonstration of genetically determined chemical markers shows that honeybees

possess a system with sufficient discriminating power for subfamily recognition.

**Gérard Arnold, Brigitte Quenet**

**Jean-Marie Cornuet, Claudine Masson**

*Laboratoire de Neurobiologie Comparée des Invertébrés, INRA-CNRS (URA 1190), BP 23, 91440 Bures-sur Yvette, France*

**Benoît De Schepper**

*Laboratoire de Chimie Agro-industrielle, ENSCT, 31077 Toulouse, France*

**Arnaud Estoup**

*Laboratoire Populations, Génétique et Evolution, CNRS, 91198 Gif-sur-Yvette, France*

**Patrick Gasqui**

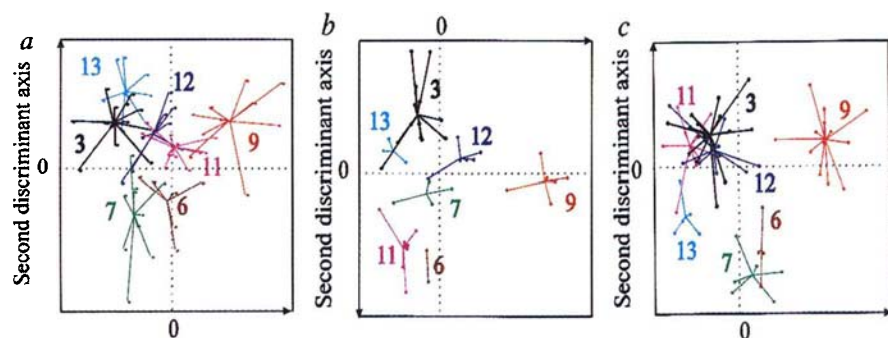
*Laboratoire de Biométrie, INRA, 78352 Jouy-en-Josas, France*

1. Estoup, A., Solignac, M. & Cornuet, J. M. *Proc. R. Soc. B* **258**, 1–7 (1994).
2. Getz, W. M. & Smith, K. B. *Nature* **302**, 147–148 (1983).
3. Moritz, R. F. A. & Hillesheim, E. *Anim. Behav.* **40**, 641–647 (1990).
4. Page, R. E., Metcalf, R. A., Metcalf, R. L., Erickson, E. H. & Lampman, R. L. *J. chem. Ecol.* **17**, 745–756 (1991).
5. McCullagh, P. & Nelder, J. A. *Generalized Linear Models* (Chapman & Hall, London, 1989).
6. Page, R. E., Robinson, G. E. & Fondrk, M. K. *Nature* **338**, 576–579 (1989).
7. Visscher, P. K. *Behav. Ecol. Sociobiol.* **18**, 453–460 (1986).
8. Oldroyd, B. P., Rinderer, T. E., Schwenke, J. R. & Buco, S. M. *Behav. Ecol. Sociobiol.* **34**, 169–173 (1994).
9. Oldroyd, B. P., Rinderer, T. E. & Buco, S. M. *Nature* **346**, 707–708 (1990).
10. Breed, M. D. *et al. Behav. Processes* **33**, 25–40 (1994).

## Fossil mesothele spiders

**SIR**—The living spiders *Liphistius* and *Heptathela* constitute the suborder Mesothelae, which is a sister group to Opisthothelae to which all other spiders belong<sup>1</sup>. Mesothelae exhibit the most primitive characteristics of all living spiders and would be expected to appear earlier in the fossil record than the oldest opisthothelae, *Rosamygale* from Triassic (240 Myr) strata<sup>2</sup>. Indeed, many Carboniferous (355–290 Myr) spiders were once referred to this suborder<sup>3</sup>. Re-examination of all available types of Carboniferous spiders reveals that some are not spiders and that none shows autapomorphies of Mesothelae. New fossils from Montceau-les-Mines, near Autun, France, however, confirm the presence of mesothelae in the late Carboniferous (around 295 Myr).

The two fossils came to light in the collections of the Natural History Museums of Autun and London (details of the new species will be published elsewhere<sup>4</sup>). The Autun specimen (*a, b* in the figure) reveals a deep, narrow sternum as wide as the labium, pro- and retromarginal cheliceral tooth rows, and two opisthosomal book-lung opercula. Paired internal structures lying above the second operculum and opening to its posterior border are similar to the supposed tracheal organs of *Heptathela*<sup>5</sup>. Posterior to the second



The distinctiveness of subfamily profiles is illustrated by a two-dimensional projection on the first plane of a factorial discriminant analysis (Statgraphics 7.0). First and second axes are linear combinations of the hydrocarbon concentrations. In the multi-dimensional space of the hydrocarbon profiles they form the plane where the projected distances between subfamilies are largest. In this analysis, performed on each set separately (*a*, isolated bees; *b*, grouped bees; *c*, hive bees), only the seven best represented subfamilies were kept. Not only were bees of the same subfamily well discriminated, but the relative positions of the subfamilies in the plane remained roughly the same in the three analyses. Each subfamily is given a different colour, and bee dots of the same subfamily are linked together through their common barycentre.



operculum, three holes represent external moulds of the left anterior lateral spinneret and the anterior median spinneret pair. Posterior to these, two more holes of approximately the same size and straddling a mid-ventral line, suggest a single pair of posterior spinnerets. Four posterior spinnerets (smaller posterior median and larger posterior lateral) are present in all spiders, except when the posterior median spinnerets are missing (some specialized opisthotheles), or a single posterior median spinneret is present (*Heptathela*)<sup>6</sup>. The posterior median spinnerets are usually small or lost, and so the relatively large posterior spinnerets present in the fossil are most probably posterior lateral spinnerets. The London specimen (*c*, *d* in the figure) shows the dorsal surface of the opisthosoma, with a series of at least ten tergites, including microtergites at the posterior end of the opisthosoma, adjacent to the anal tubercle.

The specimens are assumed to be conspecific. The animal is a spider because of the presence of spinnerets (and therefore, presumably, silk glands), a pedicel, a flexible opisthosoma with discrete tergites, carapace and sternum morphology, and lack of autapomorphies of other arachnid groups. One autapomorphy of mesotheles, a deep and narrow sternum<sup>7</sup>, can be seen; in addition, dorsal opisthosomal tergites, two ventral opercula (covering the two pairs of book-lungs), orthognath chelicerae, and fully developed anterior median spinnerets, are indicative of Mesothelae. Supposed tracheal organs are a synapomorphy for *Heptathela* and the Montceau spider (or a symplesiomorphy if homologous with amblypygid eversible sacs<sup>8</sup>). The position of the spinnerets in the fossils, neither bunched together just behind the second operculum as in modern mesotheles, nor close to the anal tubercle as in Opisthothelae, but widely spaced between the second operculum and the anal tubercle, could reflect either the start of their rearward movement towards the opisthothelate condition, or a more ancestral arrangement prior to their bunching close to the second operculum. If the posterior spinnerets in the fossils are posterior lateral spinnerets, it would represent a new configuration for Araneae, an advance over the *Heptathela* condition, and an autapomorphy for the fossil species. Another autapomorphy of the Montceau spider is biserial

cheliceral dentition (*b* in the figure; inset); only a promarginal row is present in living mesotheles<sup>9,10</sup>.

The Montceau spider is plesiomorphic with respect to Opisthothelae; the deep and narrow sternum confirms that it

belongs in Mesothelae. It is thus the first record of a fossil mesothele. The fossil species cannot be ancestral to modern mesotheles because of autapomorphies. As the sister clade to Mesothelae, Opisthothelae must have originated some time before the age of these fossils.

**Paul A. Selden**

Department of Earth Sciences,  
University of Manchester,  
Manchester M13 9PL, UK

1. Platnick, N. I. & Gertsch, W. J. *Am. Mus. Novit.* **2607**, 1–15 (1976).
2. Selden, P. A. & Gall, J.-C. *Palaeontology* **35**, 211–235 (1992).
3. Petrunkevitch, A. in *Treatise on Invertebrate Palaeontology*, Pt P. Arthropoda Vol. 2 (ed. Moore, R. C.) 42–162 (Geol. Soc. Am. & Univ. Kansas Press, Lawrence, KA, 1955).
4. Selden, P. A. *Rev. Suisse Zool.* (in the press).
5. Yoshikura, M. *Kumamoto J. Sci.* **B1**, 37–40 (1954).
6. Marples, B. J. *J. Linn. Soc. (Zool.)* **46**, 209–222, pl. 1 (1967).
7. Raven, R. J. *Bull. Am. Mus. nat. Hist.* **182**, 1–180 (1985).
8. Kaestner, A. *Invertebrate Zoology* Vol. 2 (transl. Levi, H. W. & Levi, L. R.) 123 (Wiley, New York, 1968).
9. Abraham, H. C. *Proc. zool. Soc. Lond.* **1929**, 671–677, pls 1–3 (1929).
10. Bristowe, W. S. *Proc. zool. Soc. Lond.* **1932**, 1015–1057, pls 1–6 (1932).

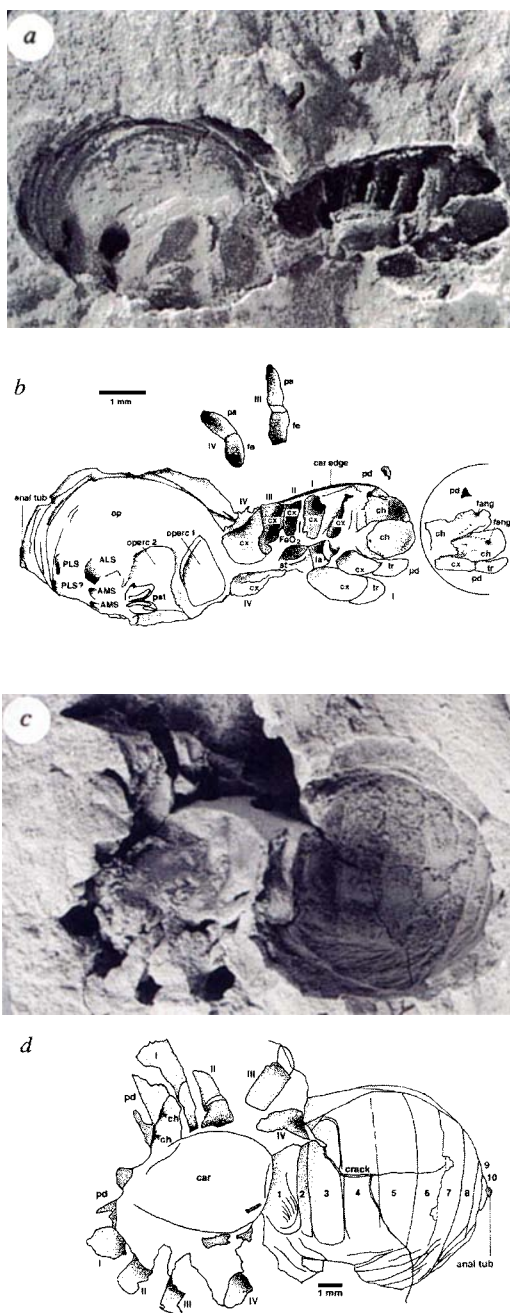
## A plant oncogene as a phosphatase

**SIR** — The plant oncogene *rolB*, from *Agrobacterium rhizogenes*, induces differentiation and growth of neoplastic roots (hairy-roots<sup>1</sup>) in dicotyledonous plants. *rolB*-transformed plant cells show an increased membrane sensitivity to<sup>2,3</sup>, and binding capacity of<sup>4</sup>, auxin, the most extensively studied plant hormone. The oncogene *rolB* may thus provide a tool for elucidating the still elusive mechanism of auxin signal perception/transduction and for shedding light on the role of this plant hormone in the control of plant growth and differentiation. So far, all attempts to clarify the biochemical activity and sub-cellular localization of the *rolB* gene product have been inconclusive. Here we show that the RolB protein overproduced in *Escherichia coli* has tyrosine phosphatase activity, and that in transformed plant cells it is localized in the plasma membrane.

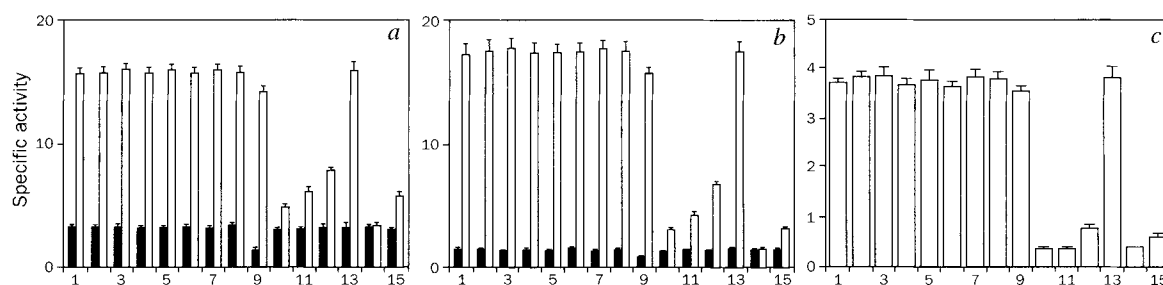
The full-length *rolB* gene was cloned in vector pE. coli W3110lac I<sup>8</sup>L8. On induction by isopropyl-β-D-thiogalactoside, the production of the RolB protein was confirmed by immunoblotting. The lysates of bacteria producing RolB (MTB4) have a fivefold higher phosphatase activity.

### Scientific Correspondence

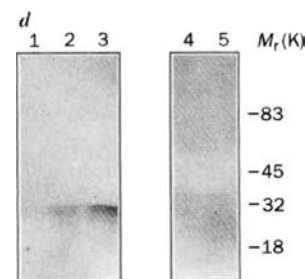
Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words.



Upper Carboniferous mesothele spider from Montceau-Mines, France. *a*, Photograph of Autun specimen 51961, whitened with  $\text{NH}_4\text{Cl}$  ( $\times 7.4$ ); *b*, explanatory drawing of *a* (inset: dorsal view of external moulds of chelicerae, to same scale); *c*, photograph of London specimen 50260, whitened with  $\text{NH}_4\text{Cl}$  ( $\times 3.3$ ); *d*, explanatory drawing of *c*. Abbreviations: 1–10, opisthosomal tergite numbers; I–IV, walking legs I–IV; ALS, anterior lateral spinneret; AMS, anterior median spinneret; anal tub, anal tubercle; car, carapace; ch, chelicera; cx, coxa; pst, paired structures; fe, femur; la, labium; operc, book-lung operculum; op, opisthosoma; pa, patella; pd, pedipalp; PLS, posterior lateral spinneret; st, sternum; te, tergite; tr, trochanter. Photos: J. Almond (*a*) and L. Anderson (*c*).



Phosphatase activities of control (W3110, closed bars) and RolB-producing (MTB4, open bars) *E. coli* lysates on: *a*, *p*-nitrophenylphosphate (pNPP); *b*, phosphotyrosine (P-Tyr); and *c*, the phospho-Tyr peptide (PPS1). Specific activity is reported as  $\mu\text{mols}$  released inorganic phosphate per h per mg protein (*a*, *b*), or nmols dephosphorylated PPS1 per h per mg protein (*c*). Columns: 1, untreated; 2, + okadaic acid; 3, + calyculin A; 4, +  $\text{Ca}^{2+}$ /calmodulin; 5, + EGTA; 6, +  $\text{Mg}^{2+}$ ; 7, + EDTA; 8, + tetramisole, levamisole; 9, + Na-tartrate; 10, + vanadate; 11, + phenylarsine oxide; 12, +  $\text{Zn}^{2+}$ ; 13, + preimmune serum; 14, + anti-RolB serum; and 15, + anti-P16 IgGs. Polyclonal anti-RolB antibodies were obtained as reported<sup>5</sup>. A peptide (P16) corresponding to amino acids 104–119 of the *A. rhizogenes* agropine RolB was synthesized by Fmoc solid-phase synthesis and anti-P16 IgGs were obtained from rabbits and purified on P16–Sephacrose columns. The phosphatase assays with pNPP or P-Tyr were performed at 30 °C at pH 7.5 and the release of inorganic phosphate measured. PPS1 corresponds to the hirudin 53–65 C-terminal fragment phosphorylated on Tyr 63. Samples were preincubated with inhibitors, stimulators or antibodies (1:500) 20 min before the substrate was added. *d*, Immunodetection of RolB in fractionated rolB-transformed (lanes 1–3) and control carrot (lanes 4, 5) cells. Lanes 1 and 4, total extract; lane 2, microsomes; lanes 3 and 5, plasma membranes. Protein markers are expressed as  $M_r \times 10^3$ . rolB transformation was carried out with a construct containing the whole *rolB* gene and its complete promoter sequence. Broken cells, after removal of the debris, constitute the 'total extract'. Microsomes were obtained, after elimination of the organelles, by resuspending the pellet of a centrifugation at 100,000g for 1 h. Plasma membranes were purified by Dextran–PEG two-phase partitioning of the microsomal fraction. Samples were separated by SDS–PAGE and blotted onto a nitrocellulose filter; the anti-P16 IgGs were used at a 1:100 dilution. The blot was developed following a common alkaline phosphatase method.



phatase activity on *p*-nitrophenylphosphate (see *a* in the figure, column 1) and 12-fold higher activity on phosphotyrosine (see *b* in the figure, column 1) than do uninduced controls. No difference was observed in the phosphoserine- and phosphothreonine-

phosphatase activity of MTB4 and control lysates (data not shown). The tyrosine phosphatase specificity of RolB was further assessed using a set of inhibitors and stimulators. None of the tested inhibitors and stimulators of Ser/Thr phosphatases

(see *a*, *b* in the figure, columns 2–7) and alkaline or acid phosphatases (columns 8, 9) affected the activity of MTB4 extracts. In contrast, the tyrosine-phosphatase inhibitors vanadate, phenylarsine oxide and zinc strongly inhibited the MTB4 phosphatase activity (see *a*, *b* in the figure, columns 10–12). The protein tyrosine phosphatase (PTPase) activity of RolB was assayed using a specific enzyme-linked immunosorbent assay (Boehringer). Only the MTB4 lysates dephosphorylated the phosphotyrosyl-peptide substrate, and the effects of inhibitors and stimulators confirmed the PTPase activity of RolB (see *c* in the figure). We also tested the effects of anti-RolB antibodies<sup>5</sup> and oligoclonal antibodies against a RolB peptide (P16). In all assays both antibodies, but not the pre-immune serum, inhibited the phosphatase activity of MTB4 extracts (see *a–c* in the figure, columns 13–15).

The PTPases isolated so far are either soluble or membrane-associated. Immunolocalization of the RolB protein in *rolB*-transformed carrot cells showed that RolB was apparent only in the plasma membrane fraction (see *d* in the figure).

In addition to ascribing a long-sought function to the product of the plant oncogene *rolB*, the PTPase activity and membrane localization of RolB reported here strongly suggest that a kinase/phosphatase cascade plays a major role in the signal transduction of the plant hormone auxin.

**Francesco Filippini**

**Valeria Rossi**

*Dipartimento di Biologia,*

**Oriano Marin**

*Dipartimento di Chimica Biologica,*  
*via Trieste, 75 - 35121 Padova, Italy*

**Maurizio Trovato**

**Paolo Costantino**

*Dipartimento di Genetica e Biologia*  
*Molecolare,*

*Università "La Sapienza",*  
*00185 Roma, Italy*

**Patrick Mark Downey**

**Fiorella Lo Schiavo**

**Mario Terzi**

*CRIBI, via Trieste, 75,*  
*35121 Padova, Italy*

1. Capone, I. et al. *Pl. molec. Biol.* **13**, 43–52 (1989).
2. Cardarelli, M. et al. *Molec. gen. Genet.* **208**, 457–463 (1987).
3. Barbier-Brygoo, H. et al. *Pl. J.* **1** (1), 83–93 (1991).
4. Filippini F., Lo Schiavo, F., Terzi, M., Costantino, P. & Trovato, M. *Pl. Cell Physiol.* **35** (5), 767–771 (1994).
5. Trovato M., Cianfriglia, M., Filetici, P., Mauro, M. L. & Costantino, P. *Gene* **87**, 139–143 (1990).

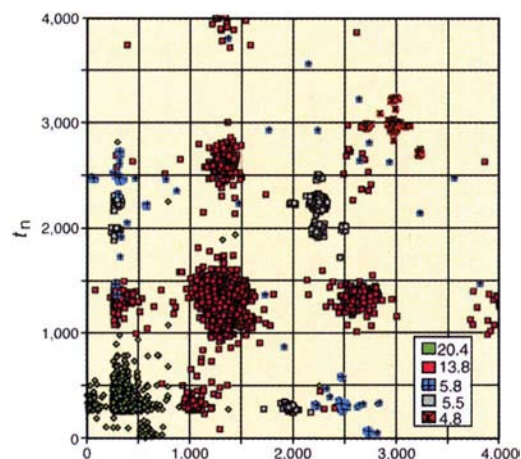
## Erratum

### Chaos in the classroom

**Justin Marston**

*Nature* **379**, 306; 1996

In this Scientific Correspondence, the figure was inadvertently printed in black and white instead of colour, in all except the US issues. It is shown here correctly.



# Nothing but the truth?

George J. Annas

**Science at the Bar: Law, Science, and Technology in America.** By Sheila Jasanoff. Harvard University Press: 1995. Pp. 283. \$29.95, £18.95.

MANY scientists, as exemplified by opinion pieces in journals such as *Science* and *Nature*, worry about the misuse of science in the courtroom, fret about judicial ignorance and activism, and expect the public hysteria and cynicism they see bred by lawsuits to stifle technological innovation. These worries extend even into the criminal courts. In October 1994, for example, Eric Lander and Bruce Budowle wrote a Commentary (*Nature* 371, 735; 1994) on the eve of the O. J. Simpson case saying that the forthcoming trial could be a "cause for rejoicing among scientists" because it would be "the most detailed course in molecular genetics ever taught to the US people". Unfortunately, they lamented, "the syllabus is being prepared by attorneys whose primary roles are as adversaries; the likely result is confusion". The purpose of their piece was magisterial: to declare that, as a matter of science, "the public needs to understand that the DNA fingerprinting controversy has been resolved". The attempt to portray science and law as separate and distinct activities, with only science being dedicated to the truth, is commonplace among scientists — and more recently among the manufacturers of products that pose dangers to consumers.

## Naivety

Sheila Jasanoff, professor of science and technology studies at Cornell University and one of America's leading scholars on science and law, thinks this science-versus-law debate is beneath contempt and beside the point. DNA fingerprinting, for example, was developed for use in court, and debates about how to declare a match and how to calculate probabilities are largely irrelevant outside the legal context. Jasanoff describes commentators such as the executive editor of the *New England Journal of Medicine* who bemoan the lack of data in litigation over silicone breast implants as espousing "naïve historiography". She states that they assume "that judges and juries could in principle have had access to 'data' in this case", pointing out that the data were produced only in the context of litigation itself and solely for this use. And to commentators, including me, who have argued that scientific matters are better left to scientific journals, and legal standards of admissibility of evidence better left to judges, she says we "fall victim to oversim-

plification" and are "historically and sociologically naïve". Her project is to use deconstruction, civic education and effectiveness as "interlinked analytic criteria" to analyse how the courts use science. For Jasanoff these criteria are "more compatible with the interactive and mutually constitutive nature of the relationship between science" and the legal process because they permit us to escape the "conceptual straightjackets" of "mainstream science" and "junk science".

In an earlier book, *The Fifth Branch: Science Advisers as Policymakers* (Harvard University Press, 1990), Jasanoff used two main metaphors to explore the history and sociology of scientific advisers and advisory committees in the US regulatory process: construction and boundaries. In her view scientific evidence in the areas of health-and-safety risk regulation, including drug safety, environmental and workplace health and safety, and regulation of carcinogens, was not so much found or discovered by scientists as constructed by them in a particular context and for a particular purpose. Both scientists and policy-makers have reasons to erect boundaries between science and policy: that is, between facts and values. When viewed in the context of history and sociology, however, these constructions can be deconstructed (taken apart) and the boundaries become illusory. Science is often manufactured only to try to address a social policy or regulatory question, and scientists and advisory committees have their own values and views of appropriate social policy that inevitably affect their conclusions. This is especially true in contentious areas where the science involved is at the frontiers and is developing hand in hand with regulation. In *Science at the Bar*, Jasanoff imports these two metaphors from the regulatory arena to the courtroom, and applies them first to science and then to technology. Each area receives four chapters and is bounded by introductory and concluding chapters. Jasanoff admits that the dividing line between science and technology "is not absolutely clear cut" and makes no real attempt to draw one. This creates some confusion in the presentation of her argument.

Scientists are almost always asked to testify in court because some technology is said to have caused some harm, or to help identify the perpetrator of a crime or an injury. Jasanoff is totally credible in her

history of appellate courts that have reviewed the decisions of administrative agencies charged with protecting the public's health and safety. In her view these appellate courts have appropriately forced the agencies to "make themselves understandable to the lay public" and thus helped both to educate the public and to democratize technical decision-making in government. I think she is also correct in concluding, on the basis of a review of legal cases involving human experimentation, use of animals in science, and peer review, that "on the plane of ideology science has little to fear from the courts". Based on their ability to deconstruct expertise, promote civic education and resolve disputes, the courts must be judged to have done a creditable job in areas involving scientific questions, even if there is room for more use of independent experts and for education of judges themselves about science, especially epidemiology.

## Values

Other conclusions, however, rest on shakier grounds. Jasanoff devotes large portions of two science chapters to medicine, as well as focusing two of her four technology chapters exclusively on medical technologies: one on technologies at the beginning of life (especially abortion technology and *in vitro* fertilization) and one on technologies at the end of life. This is understandable, but problematic. Clinical medicine is not science, and medical technology is a unique type of technology in that it is delivered in a doctor-patient relationship and can have immediate life-and-death consequences. Judges also tend to identify with physicians, as professional decision-makers, much more than they identify with perhaps any other profession, certainly much more than they identify with scientists. Judicial handling of medical technology may therefore tell us more about what judges think of physicians and their role in society than how judges deal with technology.

The two main decisions Jasanoff explores in the medical technology arena, *Roe v. Wade* and *In re Karen Ann Quinlan*, illustrate this. Both involve values at their core, but just which values remains contested. Jasanoff properly reads the decision in *Roe v. Wade* as being unable to solidify public support for a woman's right to terminate her pregnancy before the fetus is viable. But she does not point out that the language of the opinion itself focuses more on physicians' rights than women's rights. In the court's words, "The decision [*Roe*] vindicates the right of the physician to administer medical treatment according to his professional judgement. The abortion decision in all its aspects, is inherently and primarily a medical deci-



sion and the basic responsibility for it must rest with the physician." Similarly, by granting physicians legal immunity from lawsuits if they gain the assent of an "ethics committee" before removing the ventilator from a patient in a persistent vegetative state, the *Quinlan* court was trying to give physicians authority to make 'medical' decisions free of legal constraints. Jasanoff sees the 'right-to-die' courts that followed *Quinlan* as "constructing" patients (for example, as terminally ill, disease-riddled and suffering) — but what these courts really constructed was the medical profession and what judges saw as medicine's proper areas of expertise.

Jasanoff concludes that the courts have done well in the right-to-die cases by affirming the patient's ultimate right to refuse any medical treatment. Recent empirical studies, however, including the SUPPORT study (not available to her), demonstrate virtually no measurable effect on end-of-life care in US teaching hospitals when physicians are told of their patient's wishes. Her endorsement of the call by the *Quinlan* court to "delegate" authority to "ethics committees" to "micromanage" right-to-die decisions is surprising because there is no uniform definition of even what these committees are, let alone how they operate (which they virtually all do in secret), or even whether they respect the wishes of patients. Jasanoff draws an analogy between these committees and juries, but surely they act more like the scientific advisory committees she analysed so well in *The Fifth Branch*.

It is also worth mentioning that in the medical technology arena, trial courts, faced with real people in dire straits, are more likely simply to do whatever physician-experts testify is reasonable under the circumstances than are appeals courts that operate in a more abstract and rarefied environment. Jasanoff herself also often seems to be abstracted from human drama, especially in dealing with medicine: her descriptions are bloodless and fail to account for the life-changing immediacy of medical disputes that commonly reach US courtrooms. Television dramas (such as "ER" and "Chicago Hope") and nightly news broadcasts and daily newspapers give a more realistic account of the impact of medical technology on society.

This is a perceptive and elegantly written book on how science and law interact both to produce knowledge and to resolve conflict, and deserves to be read for these topics alone. The definitive archaeology of technology in the courtroom, especially medical technology, requires deeper digging. □

George J. Annas is in the Department of Health Law, Boston University Schools of Medicine and Public Health, Boston, Massachusetts 02118, USA.

## Cultivating knowledge

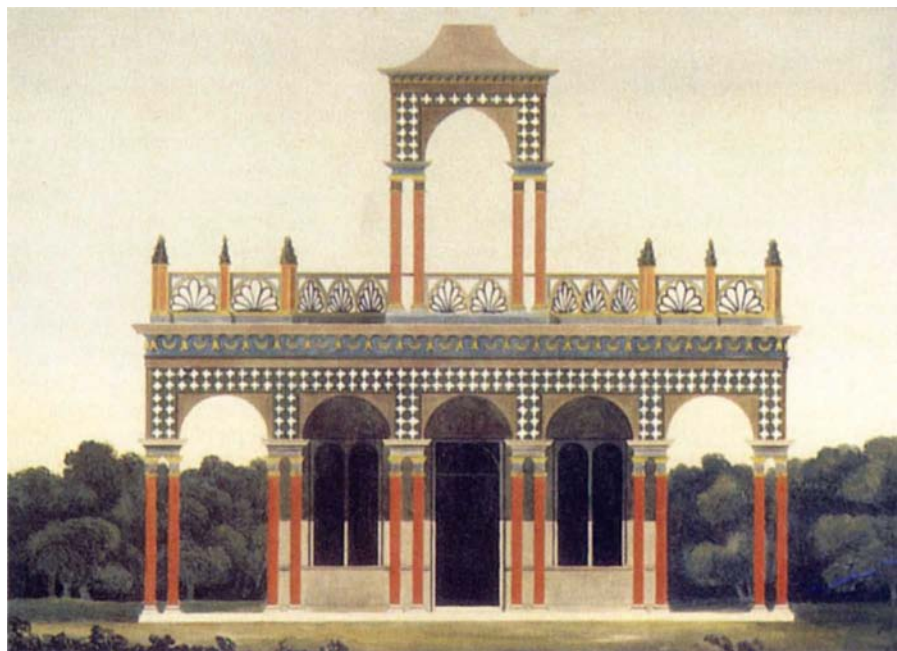
Peter H. Raven

**Kew: The History of the Royal Botanic Gardens.** By Ray Desmond. *Harvill: 1995.* Pp. 466. £25.

THE Royal Botanic Gardens at Kew have always been the very model of what a botanic garden can be in the modern world: an institution contributing to knowledge about plants, exhibiting a diverse array of species and educating a wide audience about them. Lovingly and beautifully written by Ray Desmond, long-time chief librarian and archivist at Kew,

As with all institutions, the story of Kew is the story of individuals: sometimes visionary, sometimes jealous, sometimes limited in their view of what might and should be accomplished with the assets at hand. The seminal vision was no doubt that of Sir Joseph Banks, who guided Kew's destiny from the 1770s until his death in 1820. Banks made Kew an international institution, active throughout the expanding British Empire, concerned with the flora of the world, and vital for human welfare.

Despite Banks's early leadership, Kew had by the 1830s lost its way; its establishment and growth as a government-sponsored institution were made possible only by a diverse, loosely organized network of friends, patrons and supporters throughout the country. The turning point



Chambers's Alhambra at Kew, built in 1758 and demolished in the 1820s.

this beautifully illustrated work documents the history of this great institution from its beginnings as Richmond Gardens in the early eighteenth century to today. It not only supersedes earlier histories of Kew, but also for now renders any further account of the subject superfluous.

During the eighteenth century, the story of Kew was largely one of royal patronage and the progressive development of individual gardens for various purposes, modelled along ever-changing lines of style and ingenuity. The constant revision of the basic ground plan, for instance, which involved many of England's most distinguished garden designers, including Charles Bridgeman, 'Capability' Brown and W. E. Nesfield, as well as eminent architects such as William Kent, William Chambers, James Wyatt and Decimus Burton, is a fascinating and important contribution to the history of landscape architecture.

came with John Lindley and the public commission he chaired in 1838. Plans that he drew up with his committee for the modernization of Kew were consolidated with the appointment as director of William Jackson Hooker, then professor of botany at Glasgow, in 1841. Hooker's energetic attention to the needs of the heterogeneous institution gradually converted it into the scientific force that it has remained to this day.

There have long been two competing visions of Kew: as a public park that suits the needs primarily of city-bound Londoners, but also of national and international visitors, in search of recreation; and as a scientific institution that, in expanding and disseminating knowledge, serves the needs of the entire world. Clearly, a botanic garden fulfils its functions properly when it combines all these themes.

In this respect, Kew's overseas activities

have been extraordinary. Working throughout all of the empire, its botanists and correspondents played a crucial part in the development of economic plants, transferring, for example, breadfruit from the Pacific to the West Indies, cinchona to tropical Asia and rubber from Brazil. Kew also made important contributions to the production of a series of colonial floras (descriptive plant lists) of many regions. These floras were conceived by Hooker as inexpensive volumes, readily accessible to those interested in developing the economic value of plants and understanding them better. Always at the cutting-edge, Kew is today involved in transferring the information in such floras into electronic form so that it is available throughout the world, thereby continuing a tradition pursued by its botanists and correspondents for more than 150 years.

Desmond concludes his account with a warning about the decline in biodiversity,

pointing out that a quarter of the world's 250,000 plant species may well have become extinct by the middle of next century. Such a turn of events would have been incomprehensible during the first 250 years of Kew's history. Today, however, this venerable institution is at the forefront of efforts to deal with human relationships with plants. By celebrating our dependency on plants and investigating and describing ways in which we deal with them, Kew is taking concrete steps to bring about their preservation.

This is a marvellous book, which should be read and studied by everyone interested in the history of botany, science or general history. It is an appropriate tribute to the greatness that has been, and still is, a characteristic of one of Britain's leading scientific and educational institutions. □

Peter H. Raven is at the Missouri Botanical Garden, PO Box 299, St Louis, Missouri 63166-0299, USA.

## Mycophobia and mycophilia

David L. Hawksworth

**Mushrooms: Poisons and Panaceas — A Handbook for Naturalists, Mycologists, and Physicians.** By Denis R. Benjamin. W. H. Freeman: 1995. Pp. 422. \$59.95, £54.95 (hbk); \$34.95, £24.95 (pbk).

WILD mushrooms are generally approached with suspicion in the United Kingdom, and it was perhaps inevitable that the term 'mycophobia' (later dubbed 'fungophobia') was coined by a British mycologist, W. D. Hay, in 1887. Yet Hay's countrymen are fascinated by them, and the 1,906-strong British Mycological Society is celebrating its centenary. Denis Benjamin's epilogue encapsulates why fungi merit attention: "[they] are integral to the success of life on Earth". Fiction, including Lewis Carroll's *Alice in Wonderland* (1865), H. G. Wells's *The Purple Pileus* (1895) and more recently Brian Lumley's *Fruiting Bodies* (1988), has fed on and promoted mycophobia.

The applications of mushrooms and other larger fungi in food and medicine are manifold, especially in China where *Ganoderma lucidum* is viewed as a herb of immortality, yet most of the beneficial uses claimed by indigenous peoples await objective assessment. Benjamin demonstrates that mushrooms have protein and carbohydrate concentrations falling between those of most vegetables and meat or fish; a generous serving of *Agaricus bisporus* provides about 80 kilocalories.

In commending their value as a renewable resource, Benjamin cites the collection of Pacific yew bark for the

anti-tumour compound taxol as endangering the tree; that in 1993 this same compound was found to be formed by a novel fungus living inside the tree emphasizes this message. But mushroom poisoning is the author's forte; he is a professional pathologist in Seattle and also an "amateur" mycologist tutored by Joe F. Ammirati (co-author with J. A. Traquair and P. A. Horgen of *Poisonous Mushrooms of the Northern United States and Canada*, University of Minnesota Press, 1985). Benjamin wrote this myth-debunking and eminently practical book in Edinburgh, exposed to the infectious enthusiasm of Roy Watling, another authority on poisonous mushrooms.

In discussing the nonsensical but still often used term 'toadstool', it would have been amusing to cite the legal wranglings in 1959 as to whether soup made with *Boletus edulis* should be called 'toadstool' rather than 'mushroom' soup. The prosecuting counsel asked Professor E. J. H. Corner: "What do you suppose an undergraduate, a really raw undergraduate, would expect when he buys a packet of mushroom soup?". Corner responded: "That is a very difficult question to answer... you see we never get any raw or really raw undergraduates at Cambridge". The case was dismissed.

Twenty-six central grey pages aid the rapid diagnosis and management of mushroom poisoning; one provides a key to symptoms of the eight poisoning syndromes. Critical discussions incorporate graphic case reports, including deliberate self-investigations. Most at risk are children and the "mycologically naive", and

Benjamin warns of idiosyncratic responses as well as reputed differences in toxicity between geographically separated populations. Panic attacks can also be triggered, as he experienced after a meal on suspecting he had made a wrong identification. To safeguard other mycophiles suffering such an indignity, bulleted guidelines assist the neophyte or neurotically cautious; a table of 'look-alikes' is most welcome. Adherence to the author's catechisms would preclude many unpleasant incidents, but cannot exclude risks from concentrated toxic metals and Chernobyl radionuclides.

Benjamin could have warned readers about uninvestigated and previously unknown species. Some 160 mushrooms and other fungi new to science are described from the United States each year; in the United Kingdom, mycologically the world's best studied country, the figure averages 46. Six of the 22 hallucinogenic *Psilocybe* species cited have been described in recent decades, and *Amanita smithiana*, a cause of renal failure and readily confused with the sought-after matsutake *Tricholoma mangivalve*, was named from Washington state only in 1969.

Only 36 species are photographed here in colour; *Poisonous Mushrooms* has 234. The fungi treated are primarily North American, and although many are also European, there is a mismatch. Thirty-one species names mentioned in Watling's practical *Children and Toxic Fungi* (Royal Botanic Garden, Edinburgh, 1995) are absent, as are 448 in the superlative European *A Colour Atlas of Poisonous Fungi* by A. Bresinsky and H. Besl (Wolfe, 1990).

*Wild Mushrooms and Toadstool Poisoning* by S. C. Oldridge, D. N. Pegler and B. M. Spooner (Royal Botanic Gardens,



**Better kicked than picked — the peppery, bitter-tasting *Russula emetica* can cause severe gastrointestinal upset if eaten raw.**



Kew, 1989) provides an authoritative short guide for Old World mycophiles, but anywhere between nought and four pages are devoted to poisoning in seven widely used European mushroom field guides. Benjamin's book will complement them, and the anecdotal style and liberal quotations recompense a barren day's foraging.

*Mushrooms* will not occasion a miraculous conversion to mycophily in traditionally mycophobic countries, but it promotes a rational response to mushrooms and could even save your life. □

David L. Hawksworth is at the International Mycological Institute, Bakeham Lane, Egham, Surrey TW20 9TY, UK.

## Landscapes in transition

David R. Montgomery

**The Changing Earth: Rates of Geomorphological Processes.** By Andrew Goudie. Blackwell: 1995. Pp. 302. £65, \$74.95 (hbk); £19.99, \$29.95 (pbk).

How fast do soil and rocks move around? This deceptively simple question lacks a simple answer, as processes forming the Earth's landscapes vary dramatically in nature and rate. Andrew Goudie's *The Changing Earth* presents an engaging summary of published research on rates of geomorphological processes, providing a unique resource for those concerned with the form and dynamics of the Earth's surface. It should interest professionals, researchers and advanced students in the environmental sciences.

Most of the compiled data come from the second half of the twentieth century, when development of new methods for measuring rates of geomorphological processes helped to drive the shift towards process-oriented geomorphology and away from traditional qualitative models of landscape evolution. Today, digital topographical data and increasingly sophisticated simulation models promise to revolutionize understanding of landscape evolution by better linking these two approaches. Although fulfilling this promise requires knowing the rates of the processes that build and sculpt real landscapes, compendia of these rates remain conspicuously absent.

Goudie's book provides a succinct review and synthesis, with chapters on weathering, fluvial, hillslope, aeolian, glacial, coastal and tectonic processes. The strengths of the book lie primarily in the critical review of methods for measuring geomorphological processes and the compilation of rates drawn from the literature. Each chapter contains critical

discussions of relevant field and laboratory methods, assessments particularly pertinent to those seeking an introduction to the discipline. Compilations of published studies illustrate the range of reported rates for specific processes and to some extent the relationships to factors governing these rates. Although this latter discussion provides a valuable reference, its scope defines the major weakness of the book: Goudie offers little original interpretation of the dominant controls on process rates.

Particularly effective is the use of nineteenth-century quotations to highlight the way in which our understanding of process rates and interactions resolved classic controversies, but still begs long-held questions about landscape evolution. Indeed, increasing knowledge of the interactions of processes controlling land form illustrates the naivety of many simply phrased questions about the Earth's surface. □

David R. Montgomery is in the Department of Geological Sciences, University of Washington, Seattle, Washington 98195-1310, USA.

## Under cold stone

J. L. Cloudsley-Thompson

**A Natural History of Amphibians.** By Robert C. Sebbs and Nathan W. Cohen. Princeton University Press: 1995. Pp. 316. \$29.95, £19.95.

IN recent years, both teaching and research in natural history have declined in popularity, and greater attention is now being accorded to cellular and molecular studies. Although the importance of these modern disciplines is unquestioned, it is equally important to study life at the other end of the spectrum of biological organization — knowing all there is to know about a lion's molecules and cells will not tell you why it roars. Many of the studies reported in this book provide fine examples of the first approach but, although the authors claim that it is intended for a general audience, they have not contented themselves with a mere superficial account of the subject. This is a scholarly academic review of most aspects, apart from taxonomy, palaeontology and evolution, of a class of vertebrates containing about 4,550 described living species. The bibliography alone comprises 46 pages listing nearly 1,000 carefully selected references. Not surprisingly perhaps, the emphasis is on the work of US authors — for instance the African toad *Bufo regularis* is not mentioned, although a good deal of research has been carried out on it by British zoologists — but publica-

### New in paperback

**The Origin of the Species** by Charles Darwin (2nd edn). Edited and with an introduction by Gillian Beer. Oxford University Press, £5.99. Oxford has chosen to reprint in its World's Classics series the second edition of this great work: published a mere six weeks after the first, it contains Darwin's responses — many of them climb-downs and hedges — to criticisms from his peers and the public.

**Robert Oppenheimer: Letters and Recollections** edited by Alice Kimball Smith and Charles Weiner. Stanford University Press (distributed in the United Kingdom by Cambridge University Press), \$14.95, £12.95. The letters span from Oppenheimer's graduation from Harvard in 1922 to his departure from Los Alamos in 1945, as well as recollections from his friends and colleagues.

**The Monkey Wars** by Deborah Blum. Oxford University Press, \$14.95. A balanced, wide-ranging and informative account of the issues and personalities behind the debate about the use of primates in research. Reviewed by James Jasper in *Nature* **371**, 293 (1994).

**Superforce: The Search for a Grand Unified Theory of Nature** by Paul Davies. Penguin, £7.99. A classic work of popularization, first published in 1985 and now updated and expanded.

tions emanating from other regions of the world have not been neglected.

The first 19 chapters are concerned with a variety of aspects of the biology and behaviour of extant amphibians (anatomical features, sensory structures, respiration, nutrition, voice, anti-predator defences, temperature and water regulation, home ranges and movements, migration, territorial behaviour and reproduction), whereas a final chapter is devoted to their worldwide decline in recent years. A useful addition is the indication, in parenthesis after its name, of the geographical distribution of each species mentioned in the text.

Although the authors cover much the same ground as William E. Duellman and Linda Truitt in *Biology of Amphibians* (McGraw Hill, 1986) and, to a lesser extent, Martin E. Feder and Warren W. Wurggren in their edited volume *Environmental Physiology of Amphibians* (University of Chicago Press, 1992), the new synthesis has the advantage of being more compact and up to date — knowledge of amphibian biology is today advancing rapidly. Like its predecessors, the book is well illustrated by photographs and drawings. □

J. L. Cloudsley-Thompson, emeritus professor of zoology at the University of London, is at 10 Battishill Street, London N1 1TE, UK.



# Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas

Douglas W. Burbank\*, John Leland†, Eric Fielding‡, Robert S. Anderson§, Nicholas Brozovic\*, Mary R. Reid† & Christopher Duncan||

\* Department of Earth Sciences, University of Southern California, Los Angeles, California 90089-0740, USA

† Department of Earth and Space Sciences, University of California, Los Angeles, California 90095, USA

‡ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

§ Department of Earth Sciences and Institute of Tectonics, University of California, Santa Cruz, California 95064, USA

|| Institute for the Study of the Continents, Department of Geological Sciences, Cornell University, Ithaca, New York 14853, USA

**The topography of tectonically active mountain ranges reflects a poorly understood competition between bedrock uplift and erosion. Dating of abandoned river-cut surfaces in the northwestern Himalayas reveals that the Indus river incises through the bedrock at extremely high rates (2–12 mm yr<sup>-1</sup>). In the surrounding mountains, the average angles of hillslopes are steep and essentially independent of erosion rate, suggesting control by a common threshold process. In this rapidly deforming region, an equilibrium is maintained between bedrock uplift and river incision, with landsliding allowing hillslopes to adjust efficiently to rapid river down-cutting.**

THE interactions between tectonic and erosional processes are poorly understood in most actively deforming mountain belts. Mountainous topography results from the imbalance between the upward flux of bedrock into a region due to tectonics and the outward mass flux due to denudation by either tectonic (extensional faulting) or surface (fluvial, glacial and hillslope) processes. Analysis of tectonically active landscapes therefore requires quantification of topographical characteristics, together with documentation of rates of bedrock uplift, tectonic extension and geomorphic processes. Beyond coastal areas where sea level provides a reference surface, however, it is notoriously difficult to quantify bedrock uplift<sup>1,2</sup>. Key geomorphic controls, such as rates of river incision or hillslope erosion, are often unknown, as are data constraining the relationships between rates of surface processes and topographic descriptors of the landscape, such as relief or slope angles.

Here we focus on the middle gorge of the Indus river near Nanga Parbat in northern Pakistan, in the northwestern Himalayas (Fig. 1). The north–south-trending Nanga Parbat/Haramosh axis (NPHA) has long been hypothesized to be an area experiencing rapid bedrock uplift during Quaternary times<sup>3</sup>. The suture zone (main mantle thrust, or MMT in Fig. 1) between the Indian subcontinent and the southern margin of the Kohistan–Ladakh terrane is distinctly warped around the NPHA. Bounding the western margin of the NPHA, the Raikot fault<sup>4</sup> is the only identified large-scale active fault in the region. The highest stream-gradient indices of any important Himalayan river occur where the Indus crosses the NPHA<sup>5</sup>. Apatite and zircon fission-track ages<sup>6,7</sup> and numerous radiometric dates<sup>8–11</sup> clearly indicate that the NPHA (Fig. 1) has experienced rapid cooling, which has been interpreted to result from high bedrock uplift and denudation rates<sup>6,7,10,12</sup>. The strong contrasts in radiometric ages between the NPHA and adjacent regions make this an obvious site in which to explore the reaction of topography to contrasts in tectonic deformation rates.

We quantify here key interactions between tectonic and geomorphic processes along the Indus. On the basis of cosmogenic radionuclide exposure ages of abandoned river-cut terraces, we document some of the highest known rates of bedrock incision (2–12 mm yr<sup>-1</sup>). We calculate rates of long-term denudation from

published bedrock cooling ages, and assess bedrock uplift by combining incision rates with reasonable schemes for the evolution of the longitudinal profile of the Indus. The distribution of hillslope angles is both remarkably similar among contrasting regions and apparently independent of rates of denudation or river incision. Steep mean slope angles suggest a landscape poised near the threshold for landsliding involving the bedrock. Down-cutting by large rivers roughly balances regionally variable rates of bedrock uplift. Throughout the reach of the Indus that we have studied, the rate of bedrock incision exceeds 2 mm yr<sup>-1</sup> and drives down the toes of the adjacent slopes so rapidly that they adjust most efficiently through landsliding.

## Bedrock incision rates

When a river cuts laterally into a bedrock valley wall, it carves a bench, or strath. Subsequent vertical incision causes abandonment of the strath. If the time since abandonment and the height of the strath above the valley bottom are known, a mean rate of river incision can be determined. Numerous straths are remarkably well preserved at heights ranging from 10 to 400 m or more above the Indus. Cut into metamorphic and igneous bedrock and lacking downstream continuity, many straths appear to be nearly pristine surfaces, 10 to 200 m wide, whose characteristics mimic those of river-worn surfaces along the modern channel. Straths that have experienced little erosional modification since abandonment are identified by streamlined bedrock knobs protruding 1–10 m above the mean strath surface, which is also sculpted, potholed and polished. Unlike straths in many other areas, alluvial mantles either were never present or were swept off many Indus straths. Although some straths have been buried by ancient landslides, such straths can usually be identified and were not sampled.

We collected bedrock samples near the outer edge of individual straths along the Indus between the Raikot fault and the Skardu basin (Fig. 1) for exposure-age dating using cosmogenic <sup>10</sup>Be and <sup>26</sup>Al. To take advantage of known <sup>10</sup>Be and <sup>26</sup>Al production rates in quartz<sup>13</sup>, quartz-rich samples were dated<sup>14</sup>. All <sup>10</sup>Be/<sup>9</sup>Be and <sup>26</sup>Al/<sup>27</sup>Al measurements were made by accelerator mass spectrometry at Lawrence Livermore National Laboratory. Reported ages (Table 1) are corrected for altitude, latitude and exposure geometry<sup>13,15</sup>. Cited errors (2σ) include propagation of all

measurement uncertainties; production-rate uncertainties are excluded from the error analysis. All measured  $^{26}\text{Al}$  ages (five samples) were concordant with  $^{10}\text{Be}$  ages on the same sample (Table 1). This concordance and the fact that the measured ratios of  $^{26}\text{Al}/^{10}\text{Be}$  are equal to the known production-rate ratios of these isotopes suggest that these surfaces have not experienced significant erosion or burial<sup>16</sup> and reinforce our interpretation of the sampled bedrock terraces as pristine strath surfaces. Four straths yielded consistent ages from two separate sites tens of metres apart (straths 1, 3, 4 and 7), whereas two surfaces (straths 2 and 6) have contrasting ages from separate sites (Table 1). Inconsistent ages can result from cosmic-ray screening of a site by an overlying boulder or from sampling a surface that either was more recently exposed because of spalling or which stood significantly above adjacent eroding surfaces. We have little field evidence for any of these explanations. In each case of mismatched ages on a single strath, one age is consistent with another dated strath nearby, and we have used this age in our interpretations.

Straths ranging from 65 to 410 m above the height of the Indus at low discharge yielded exposure ages spanning a range from about 4 to 68 kyr (Table 1). Making the assumption that the interval of strath formation was short in comparison to the time since abandonment by the Indus, we combined the mean exposure age of each strath and its height above the river to yield an average bedrock incision rate. High-water lines from annual floods are typically present 10 m or more above the low-stage discharges that prevailed during our sampling. Exposure ages on rock surfaces 1–5 m above the low-stage Indus yield ages indistinguishable from our process blank and indicate active erosion of these surfaces. For incision calculations, we assume that active river erosion at timescales of several decades occurs up to 20 m above low-water levels. Thus, the calculated rates are conservative, minimum rates. At the only site where we have dated vertically juxtaposed straths (straths 5 and 6), they yield consistent incision rates. Indus incision rates (Table 1) range from  $\sim 2\text{ mm yr}^{-1}$  to  $12\text{ mm yr}^{-1}$  between the Skardu basin and the Raikot fault (Fig. 2). These rates are some of the highest sustained bedrock incision rates that have been documented anywhere in the world. There is, moreover, a distinct spatial pattern of incision rates. At all sampled localities within and adjacent to the NPHA, incision rates are high, displaying maximum rates of  $8\text{--}12\text{ mm yr}^{-1}$  (Fig. 2).

Fifty kilometres to the east, rates seem to diminish slightly (to  $5\text{--}7\text{ mm yr}^{-1}$ : strath 3), whereas less than 10 km farther east, rates decrease abruptly to  $2\text{--}3\text{ mm yr}^{-1}$ . Here, across a horizontal distance of a few kilometres there is a 3-fold change in bedrock incision rates (Fig. 2).

### Long-term denudation

Whereas cosmogenic radionuclide exposure ages define short-term incision rates, they do not reveal the persistence or variability of these rates over longer time intervals. Long-term denudation rates can be estimated from radiometric cooling ages, if that cooling is the result of the upward motion of rock with respect to the Earth's surface ('bedrock exhumation'). The cooling age represents the time since a given mineral in a rock sample was at or above the mineral closure or annealing temperature. With a known or assumed geothermal gradient of increasing temperature with depth, one can estimate a depth below the surface at which the sample had to be for that temperature. The approximate long-term denudation rate is then the depth or thickness of rocks removed divided by the cooling age. Conservative estimates of annealing temperatures ( $\sim 120$  and  $210^\circ\text{C}$  for rapidly cooled apatite and zircon, respectively<sup>7,17,18</sup>) are used here, because they minimize the calculated denudation rate. Rapid cooling rates indicated by many young ages may warrant both higher annealing temperatures ( $150$  and  $240^\circ\text{C}$ )<sup>12,19</sup> and faster estimated denudation rates.

Past geothermal gradients and any temporal changes in those gradients are unknown. In nearby, but more slowly denuding Kohistan (Fig. 1), an early Miocene gradient of  $\sim 40^\circ\text{C km}^{-1}$  is definable on the basis of fission-track data.<sup>7</sup> An appropriate upper limit to the geothermal gradient of  $\sim 60^\circ\text{C km}^{-1}$  in areas of very rapid Pleistocene denudation<sup>6</sup> is suggested by pressure–temperature data and co-existing  $^{40}\text{Ar}/^{39}\text{Ar}$  ages ( $\sim 1\text{ Myr}$  on biotite) from high elevations on Nanga Parbat<sup>10</sup>. Given calculated denudation rates of  $\geq 1\text{ mm yr}^{-1}$  during the past 2 Myr (Fig. 2) in most of the study area, a likely lower limit of  $35^\circ\text{C km}^{-1}$  would be expected to result from advection and compression of isotherms in response to denudation<sup>20</sup>. We therefore use gradients of  $35^\circ\text{C km}^{-1}$  and  $60^\circ\text{C km}^{-1}$  to place probable bounds on long-term denudation rates.

The apatite fission-track dates<sup>6,7</sup> yield bedrock cooling rates ranging from  $>250^\circ\text{C Myr}^{-1}$  in the NPHA to  $<15^\circ\text{C Myr}^{-1}$  near Skardu. Comparisons of long-term denudation rates based on these dates with incision rates determined using cosmogenic nuclides for nearby strath surfaces (Fig. 2) indicate, first, that high denudation rates are associated with high incision rates across the NPHA; second, that the eastern limits of the zones of high denudation and of high bedrock incision rates are marked by abrupt, 2-to-4-fold decreases in calculated rates; third, given the uncertainties in past geothermal gradients, that long-term denudation rates during the past 0.5–1 Myr across the NPHA are strikingly similar to late Pleistocene incision rates; and fourth, that the zone of rapid bedrock incision seems to be broader than the zone of rapid long-term denudation. The greater breadth of the zone of rapid incision (Fig. 2) creates a mismatch of rates for  $\sim 50\text{ km}$  upstream of the NPHA and suggests that denudation rates in this region may have accelerated during latest Quaternary times. Overall, denudation rates similar to those defined by the incision data are likely to have persisted for 0.5 Myr or more within the NPHA, during which time  $\sim 2\text{--}4\text{ km}$  of bedrock would have been removed from the NPHA itself. Whereas such denudation rates have been estimated previously<sup>7,10,12</sup>, we are now able to compare them directly with bedrock incision rates for the same geographical localities.

Cooling rates based on zircon fission-track dates are only half as rapid as those determined from apatite dates on the same or nearby samples<sup>6,7</sup>, except in the Skardu basin, where rates are uniformly slow ( $<15^\circ\text{C Myr}^{-1}$ ). This difference in cooling rates could result from accelerating denudation rates or Pleistocene steepening of the geothermal gradient. Some combination of the

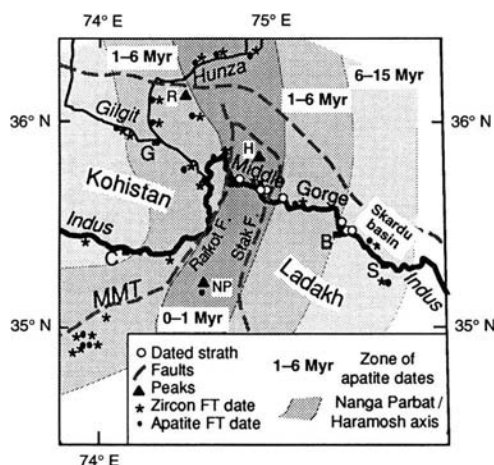


FIG. 1 Map of the Nanga Parbat region of the Indus river catchment showing main tributary rivers, dated straths in the middle gorge of the Indus, apatite and zircon fission-track (FT) dates, zones of contrasting apatite fission-track dates (indicated by different shading), and main faults (F). The youngest apatite dates correspond with the Nanga Parbat/Haramosh axis (darkest shading). B, Bashu; C, Chilas; G, Gilgit; H, Haramosh; MMT, main mantle thrust; NP, Nanga Parbat; R, Rakaposhi; S, Skardu.

TABLE 1  $^{10}\text{Be}$  and  $^{26}\text{Al}$  exposure ages for strath terraces in the Middle Gorge of the Indus river

| Strath no. | Sample no. | Distance upstream (km) | Altitude (m) | Height above river (m) | $^{10}\text{Be}$ concentration ( $10^6$ atoms per g qz) | $^{10}\text{Be}$ age (kyr) | $^{26}\text{Al}$ concentration ( $10^6$ atoms per g qz) | $^{26}\text{Al}$ age (kyr) | Averaged strath age (kyr) | Incision rate ( $\text{mm yr}^{-1}$ )* |
|------------|------------|------------------------|--------------|------------------------|---------------------------------------------------------|----------------------------|---------------------------------------------------------|----------------------------|---------------------------|----------------------------------------|
| 1          | 93-20      | 133                    | 2,235        | 185                    | $1.43 \pm 0.16$                                         | $60 \pm 7$                 | n.m.                                                    | n.m.                       | $61 \pm 6$                | $2.7 \pm 0.4$                          |
|            | 93-22      | 133                    | 2,233        | 183                    | $1.53 \pm 0.42$                                         | $67 \pm 19$                | n.m.                                                    | n.m.                       |                           |                                        |
| 2          | 93-23      | 131                    | 2,198        | 148                    | $1.49 \pm 0.14$                                         | $65 \pm 6$                 | $9.24 \pm 0.70$                                         | $68 \pm 5$                 | $67 \pm 4$                | $1.9 \pm 0.2$                          |
|            | 93-25      | 131                    | 2,195        | 145                    | $0.75 \pm 0.05$                                         | $(32 \pm 2)$               | n.m.                                                    | n.m.                       |                           |                                        |
| 3          | 93-17      | 124                    | 2,082        | 65                     | $0.15 \pm 0.01$                                         | $7 \pm 1$                  | n.m.                                                    | n.m.                       | $7 \pm 1$                 | $6.5 \pm 1.8$                          |
|            | 93-19      | 124                    | 2,082        | 65                     | $0.17 \pm 0.04$                                         | $8 \pm 2$                  | $0.92 \pm 0.28$                                         | $7 \pm 2$                  |                           |                                        |
| 4          | 93-29      | 81                     | 1,850        | 70                     | $0.13 \pm 0.03$                                         | $7 \pm 2$                  | n.m.                                                    | n.m.                       | $6 \pm 1$                 | $7.8 \pm 2.6$                          |
|            | 93-30      | 81                     | 1,840        | 65                     | $0.08 \pm 0.03$                                         | $4 \pm 2$                  | n.m.                                                    | n.m.                       |                           |                                        |
| 5          | 95-31      | 69                     | 2,150        | 410                    | $0.76 \pm 0.06$                                         | $32 \pm 2$                 | $4.50 \pm 0.65$                                         | $32 \pm 5$                 | $32 \pm 2$                | $12.1 \pm 0.9$                         |
| 6          | 93-96      | 68                     | 1,883        | 155                    | $0.24 \pm 0.13$                                         | $12 \pm 7$                 | $1.70 \pm 0.36$                                         | $15 \pm 3$                 | $14 \pm 3$                | $9.6 \pm 2.3$                          |
|            | 93-38      | 68                     | 1,882        | 154                    | $0.55 \pm 0.05$                                         | $(29 \pm 3)$               | $2.93 \pm 0.32$                                         | $(25 \pm 3)$               |                           |                                        |
| 7          | 93-39      | 52                     | 1,675        | 80                     | $0.12 \pm 0.02$                                         | $8 \pm 1$                  | n.m.                                                    | n.m.                       | $7 \pm 1$                 | $8.9 \pm 2.2$                          |
|            | 93-40      | 52                     | 1,675        | 80                     | $0.80 \pm 0.02$                                         | $5 \pm 1$                  | n.m.                                                    | n.m.                       |                           |                                        |

All uncertainties are  $2\sigma$  values. Values listed in parentheses are not used to calculate average strath age. n.m., not measured.

\* Rate is calculated assuming active erosion up to 20 m above low-discharge levels of the Indus. Incision rate error is calculated using uncertainty in average age and a 5-m uncertainty in elevation.

two seems likely. Irrespective of possible changes in the geothermal gradient, when compared with the apatite and incision data, the zircons in the NPHA also yield the highest rates of denudation and comparable contrasts of high-to-low rates across the eastern NPHA (Fig. 2). Even if the geothermal gradient increased from 35 to  $60^\circ\text{C km}^{-1}$  at all sites between the early

and middle Pleistocene, there is evidence for modest (20–50%) acceleration of denudation rates at this time. Comparisons of rates based on apatite versus strath ages suggest the accelerating trend has persisted into late Pleistocene times (Fig. 2).

### Calculation of bedrock uplift

The distinctive gradient in incision rates within the upper part of the middle Indus gorge (straths 1–3; Fig. 2) cannot be attributed to differences in erodability of the bedrock within the NPHA of the Ladakh terrane (Fig. 1), because both comprise highly resistant igneous and metamorphic rocks. Rather, contrasting rates seem to be driven by differential bedrock uplift near Bashu (Fig. 3a). We attempt here to constrain both differential uplift and absolute bedrock-uplift rates.

If we assume that incision rates defined at each site have persisted for at least 65 kyr (the approximate age of the oldest dated straths), then the height above the present Indus of a strath formed 65 kyr ago can be predicted by extrapolating short-term incision rates back through time. Thus, for the 32-kyr-old strath (strath 5) which is currently 410 m above the Indus, a 65-kyr-old strath in that location would be 830 m above the modern river. By connecting the projected positions of these straths, the deformed longitudinal profile of the Indus at  $\sim 65$  kyr can be reconstructed (Fig. 3a). The 32-kyr-old strath defines the lowest permissible altitude for this profile within the NPHA. Comparison of the reconstructed and modern Indus profiles reveals pronounced differential bedrock uplift near Bashu, downstream divergence of the longitudinal profiles, and low-gradient reconstructed 65-kyr profile, whose western half is 500–800 m above the modern Indus.

The far more difficult task of defining bedrock uplift with respect to the geoid<sup>2</sup> can be examined initially near Bashu (Fig. 3a). Given the young radiometric ages nearby<sup>7,8,12</sup>, it is highly unlikely that bedrock is subsiding with respect to the geoid. We assume conservatively that there has been no bedrock uplift in the area of the 65-kyr strath upstream of Bashu (that is, bedrock here was fixed with respect to the geoid) and that the Indus has incised steadily after abandoning the 65-kyr strath (Fig. 3b). This permits us to estimate the height of the Indus at  $\sim 7$  kyr upstream of Bashu and to compare it with the surface at the same age 7–9 km downstream (strath 3). The river gradient at 7 kyr is unknown, but it certainly was more than  $0^\circ$  and probably was no steeper than the modern gradient. The altitudinal offset of the 7-kyr strath with respect to these two gradient limits defines bedrock uplift ranging from 2 to  $5 \text{ mm yr}^{-1}$  for the downstream region (Fig. 3b). It seems unlikely, however, that the gradient of the Indus was negligible throughout this reach, because that requires improbably steep gradients farther downstream. If both the elevation and gradient of the Indus River at 7-kyr were about the same as today, then

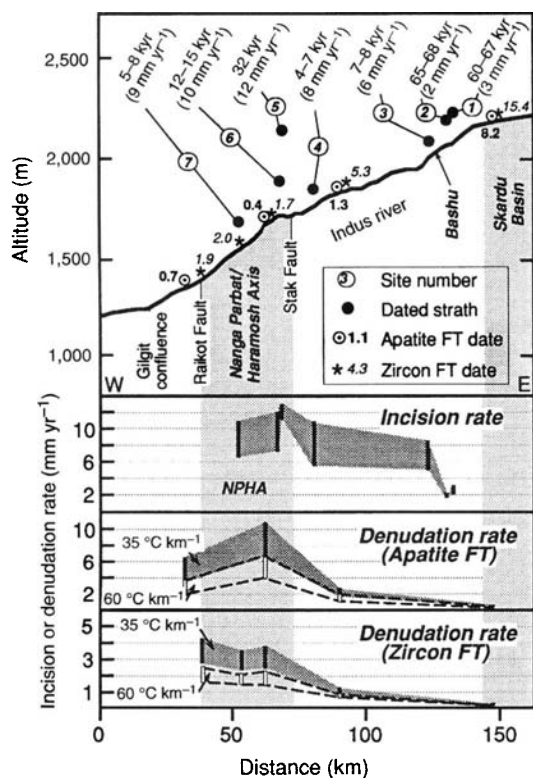


FIG. 2 Longitudinal profile of the Indus river from near its confluence with the Gilgit river, across the NPHA and middle gorge to the Skardu basin. The locations, height above the Indus, ages, incision rates and numbers of dated straths are shown along with the location and ages of spatially associated fission-track dates<sup>7</sup>. Incision rates (with  $2\sigma$  uncertainties) are highest in the NPHA (second panel). Between straths 2 and 3 near Bashu, there is a roughly 3-fold increase in incision rates. Across the eastern edge of the NPHA (corresponding roughly to the Stak fault zone), a comparable change occurs in denudation rates (lower panels) derived from apatite and zircon fission-track ages. Ranges of denudation rates ( $2\sigma$ ) are depicted for  $35^\circ\text{C km}^{-1}$  and  $60^\circ\text{C km}^{-1}$  geothermal gradients. Note change in vertical scale for zircon data.



incision by the Indus is keeping pace with bedrock uplift. This latter picture implies that the longitudinal profile of the Indus in this region could be regarded as a fixed reference frame against which the vertical movement of bedrock could be measured.

Fluvial incision into bedrock is driven by specific stream power<sup>21–23</sup>, which is proportional to the product of slope and discharge and is inversely proportional to river width. Because the middle gorge has a catchment area of  $>10^5 \text{ km}^2$  and is  $>600 \text{ km}$  from the Indus headwaters, there is considerable discharge through it (summer flows are typically  $>3,000 \text{ m}^3 \text{ s}^{-1}$ ). Moreover, no significant tributaries join the Indus in the middle gorge. The variation in the Indus gradient throughout this reach, therefore, dictates the pattern of stream power. That the Indus currently displays its maximum gradient (Fig. 3c) in and around the NPHA<sup>5</sup> implies that the longitudinal profile of the Indus may be in rough equilibrium with the bedrock uplift pattern.

Although a fixed position with respect to the geoid is probably not realistic for many rivers in actively deforming areas, we suggest that it is appropriate here during the past 0.5–1 Myr. Given the similar magnitude within the NPHA of both denudation rates and incision rates (Fig. 2), many kilometres of rock must have been eroded from the NPHA during the Quaternary. Any sustained mismatch between the rates of incision and bedrock uplift would result in considerable perturbations in the longitudinal profile of the Indus (Fig. 3c). It is, therefore, important to assess the evidence for any such perturbations.

If incision rates exceed bedrock uplift rates in the NPHA, the Indus will lower its elevation within the NPHA and necessarily lower its gradient below this reach. Unless the Indus also becomes narrower (the opposite of what is observed in the gentler reaches), the lowered gradient generates a negative feedback which decreases both the specific stream power and the rate of incision until a new balance is attained between gradient, incision rate and bedrock uplift rate. Given that the study area is  $>1,200 \text{ km}$  from

the mouth of the Indus, and that the river traverses 400 km of mountainous terrain after crossing the NPHA, the Indus simply cannot lower its downstream gradient too much. For example, if present incision rates were sustained for 100 kyr in the face of no bedrock uplift (Fig. 3c), the Indus would be nearly below sea level in the study area.

If, on the other hand, rates of bedrock uplift were greater than rates of incision across a zone of rapid differential uplift, the slope of the Indus would decrease with time across the uplift. The gradient downstream of the uplifting region, however, would increase, promoting accelerated incision. If the mismatch across the uplift persisted, the river could become tectonically dammed and/or have its flow reversed. To maintain an oceanward gradient, an upstream reach of the river must aggrade and/or be uplifted at a rate equal to the difference between uplift and incision downstream. A rapid pulse of aggradation did occur in the Skardu basin, immediately above the middle gorge, during the early Pleistocene<sup>24,25</sup> and may reflect a transient response to such a mismatch of incision rates with newly accelerated bedrock uplift rates<sup>26</sup>. There is, however, little evidence for extensive aggradation along the upper reaches of the Indus, where calculated denudation rates<sup>27</sup> are  $\sim 40$  times less than those in the NPHA. We cannot eliminate the possibility that the entire Indus drainage upstream of the NPHA experienced bedrock uplift during a part of the Quaternary at a rate nearly equal to the difference between uplift and incision rates at the NPHA. But given the present upstream elevation of the Indus ( $\sim 3,000 \text{ m}$  near Leh,  $>300 \text{ km}$  upstream) and its large distance from the ocean ( $\sim 1,500 \text{ km}$  at Leh), the low Neogene denudation rates described along its upstream course<sup>7,27</sup>, and the likely rates of bedrock incision as determined from stream power calibrated against our incision data<sup>26</sup>, it is unlikely that the elevation of the Indus upstream of the study area has increased by  $>1 \text{ km}$  during the past 1 Myr. Because this represents  $<25\%$  of the documented denudation across the NPHA, we conclude that

FIG. 3 a, Reconstruction of strath geometries along the middle Indus gorge. Extrapolation of site-specific incision rates (assumed constant during the past 65 kyr) define expected present-day altitudinal positions of 65-kyr-old straths (dashed line, with  $2\sigma$  error). Strong warping of reconstructed Indus course at 65 kyr ago (stippled area) illustrates clear differential uplift of the downstream reach. Note that if no bedrock uplift is assumed, the Indus in the NPHA region would have been  $\sim 1 \text{ km}$  above its present elevation less than 0.1 Myr ago. This would require very steep downstream gradients and improbable upstream river profiles. b, Two interpretations of bedrock uplift near Bashu, assuming no bedrock uplift further than 131 km upstream (strath 2) and constant incision between 65 and 7 kyr ago. The downstream strath at 7 kyr (strath 3) is altitudinally above the calculated river position upstream at 131 km. The amount and rate of bedrock uplift of strath 3 depends on the assumed gradient downstream of 131 km. A minimum bedrock uplift rate of  $2 \text{ mm yr}^{-1}$  results from assuming an improbable, flat gradient. If the gradient were parallel to the modern, the bedrock uplift rate would be  $\geq 5 \text{ mm yr}^{-1}$ . c, Comparison of the longitudinal profile of the modern Indus with profiles that would be predicted if there were no bedrock uplift, but if present incision rates persisted into the past or future. The steepest gradient of the modern Indus occurs in and near the NPHA, where incision rates are highest. At 65 kyr ago, there would have been a very gentle gradient across the NPHA and a much steeper downstream gradient than at present without bedrock uplift. Additional extrapolation to 130 kyr ago yields an even steeper downstream gradient (through an area with relatively old cooling ages; Fig. 1) and predicts a reversal of dip of the river profile across the NPHA. This is a highly unlikely geometry for the river at that time, but is the expected configuration for straths of that age, if there were no bedrock uplift. If present incision rates were to continue 65 kyr into the future without bedrock uplift, the river would be near sea level in the NPHA: clearly another improbable geometry. Note that, if the course of the Indus upstream of Skardu has been raised 1 km during the Quaternary, this would require a gentle (and unlikely) gradient across the NPHA at 1 Myr ago. These examples indicate that bedrock uplift and incision must be nearly in balance across the middle Indus gorge.

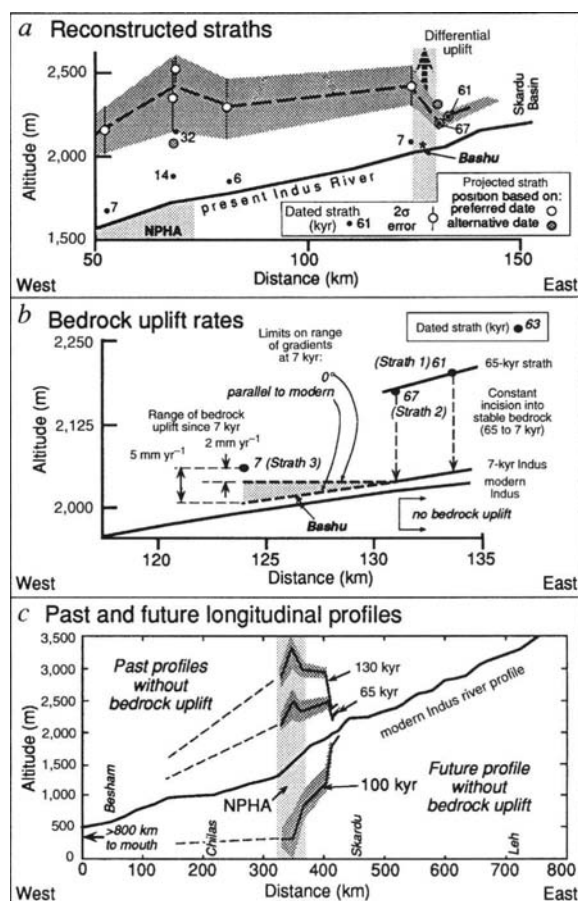
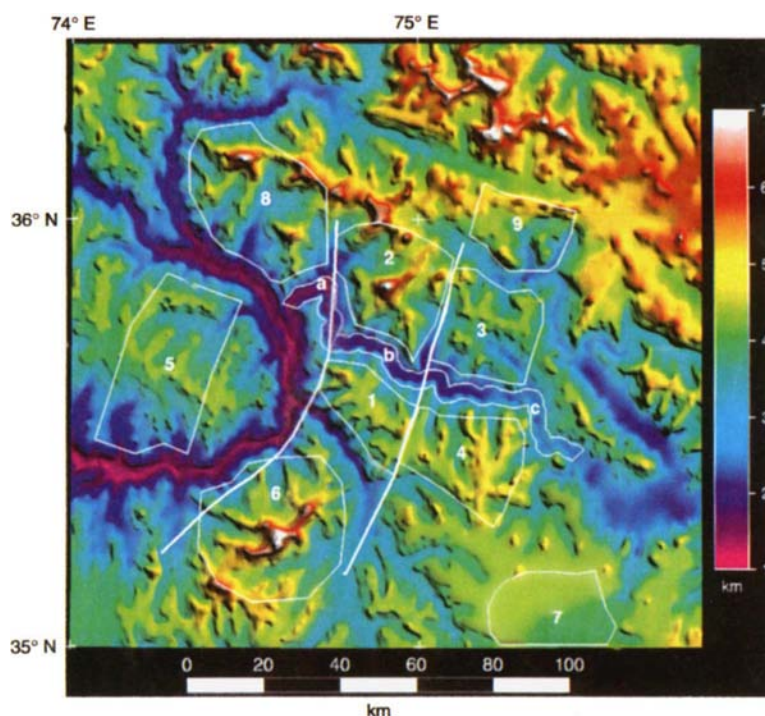


FIG. 4 Digital topography of the Nanga Parbat region of northern Pakistan, derived from public-domain data on the *Digital Chart of the World*, which has  $\sim 1\%$  of the resolution of the 3-arcsecond data used in the topographic analyses. Altitudes range from 1 to  $>8$  km. Areas analysed for slope distributions (Fig. 5) are either 4-km-wide swaths along the Indus river or extensive mountainous (areas 1, 2, 3, 6, 8 and 9) or plateau-like regions (areas 4, 5 and 7) next to it, shown as white polygons. Nanga Parbat is in region 6, Haramosh in region 2, Rakaposhi in region 8, the Gamugah surface in region 5, and the Deosai Plateau in region 7. The middle gorge of the Indus flows from swath **c** through swath **a**. The approximate traces of the Raikot and Stak faults are shown (thick white lines).



bedrock uplift and incision rates have been nearly balanced across the NPHA during much of the past 1 Myr, that the longitudinal profile of the Indus is roughly fixed with respect to the geoid, and that incision rates therefore provide a reasonable proxy for bedrock uplift rates ( $\pm 25\%$ ) during the Quaternary in the study area.

### Responses to differential uplift

It has often been argued that high rates of bedrock uplift and denudation should be correlated with steep slope angles<sup>28</sup>, because hillslope erosion rates are strongly slope-dependent<sup>23,29,30</sup>. In the Nanga Parbat region, excellent opportunities exist to test these concepts at large scales, because the incision and radiometric data define regions of sharply contrasting rates of denudation (Fig. 2), and because digital topographic data permit quantification of regional topographic characteristics.

We have analysed a digital elevation model (DEM) with 3-arcsecond ( $\sim 90$ -m) grid spacing<sup>31</sup>. Slope angles are defined using a  $4 \times 4$ -point window within which the slope of a best-fitting plane is calculated to the nearest  $0.1^\circ$ . The  $\sim 270$  m size of the measurement window causes slopes to be underestimated, especially in areas where hillslopes have substantial curvature, such as across sharp ridges or valley bottoms. For each area analysed,  $\sim 10,000$ – $40,000$  slope determinations are collected into a histogram. The resultant slope distributions are sorted into  $1^\circ$  bins and then normalized by area to permit comparison between unequally sized areas.

We have analysed six mountainous regions that sample the contrasting zones of denudation, three plateau-like regions, and three 4-km-wide swaths centred along the Indus river where we have incision or denudation data (Fig. 4; note that Fig. 4 does not show the 3-arcsecond data). We analysed the mountains and the Indus valley separately in case large river valleys strongly perturb slope distributions. The summits of Nanga Parbat (8,126 m) and Haramosh (7,458 m) occur in the mountainous regions analysed within the NPHA. Rakaposhi (7,788 m) lies to the northwest of the NPHA (Figs 1 and 4) near the margins of a region whose fission-track dates<sup>7</sup> suggest denudation rates similar to the relatively low rates along the Indus upstream of the NPHA. No comparably high summits lie in the selected regions to the east of the NPHA. The plateau-like areas include the central Deosai

Plateau, a mountainous area that encompasses one margin of the Deosai and the partially dissected 'Gamugah' surface (Fig. 4).

Remarkably, the distributions of slopes are essentially indistinguishable among different mountainous regions, despite significant differences in denudation rates (Fig. 5a). In each region, most slopes fall between  $20^\circ$  and  $45^\circ$ , and the mean slope angle is  $\sim 32 \pm 2^\circ$  (for 270-m windows). Some regions containing large valley glaciers, which typically have low gradients, have a secondary peak of low slope angles ( $\sim 5^\circ$ ). The swaths along the Indus show similarly high mean slope angles (Fig. 5b). In fact, the highest mean slope angle ( $38^\circ$ ) among all areas analysed occurs along the largely unglaciated reach of the Indus with the highest incision and denudation rates. Whereas the central Deosai Plateau displays low mean slopes ( $\sim 11^\circ$ ), more heavily dissected plateau remnants (Fig. 5c) show relatively high mean slopes ( $\sim 26^\circ$ ). There are few radiometric ages to constrain directly the cooling or denudation histories of these plateaux. Most dates along their margins indicate cooling rates of  $10$ – $20^\circ\text{C Myr}^{-1}$  (ref. 7), about 10 times slower than in the NPHA.

Three key conclusions can be drawn from the slope analysis. First, outside the plateaux, slope angles are largely independent of, rather than correlated with<sup>28</sup>, denudation or bedrock-uplift rates. Nor are they correlated with predominantly glaciated (mountain) zones rather than non-glaciated (valley-bottom) zones. Second, uniformly high mean slope angles among all areas (other than the plateaux) suggest that a common threshold controls them. Third, the data do not support at least one proposal for hillslope response to accelerated denudation: 'slope replacement' through the lengthening of steeper slopes at the expense of gentler slopes<sup>28,32</sup>. The histograms measure the area of each slope category, and, when one compares less rapidly with more rapidly denuding areas, the histograms do not reveal any significant shift in slope distributions due to substitution of steeper for gentler slopes.

DEM analysis and field observations indicate that bedrock-involved landsliding<sup>33</sup> is the dominant denudational process in these rapidly denuding mountains. It is also important that rivers of this region are generally underloaded with respect to sediment. They successfully remove detritus supplied to them by adjacent hillsides, and throughout most of the area, they are flowing on or near bedrock. The absence of significant quantities of stored

FIG. 5 a, Slope distributions in mountainous regions near Nanga Parbat, excluding plateaux. Each numbered histogram corresponds to a region defined on the DEM (Fig. 4). For comparison, all histograms are normalized by area. There are no significant differences between areas with high denudation rates in the NPHA (curves 1, 2 and 6) and areas with much lower denudation rates in adjacent regions. Although underestimated by the 90-m DEM, the mean slope angle for all areas is  $32 \pm 2^\circ$ . b, Slope distributions for swaths  $\sim 4$ -km wide along the Middle Indus gorge. Highest mean slope angles ( $38^\circ$ ) occur in the region of most rapid incision (area b, Fig. 4) which also displays a smaller proportion of low slope angles than any of the other studied areas, possibly because of absence of significant glaciation in the valley bottom. Compared with the zone of the most rapid incision, lower slope angles are more prevalent in the downstream and upstream zones (a and c, respectively, Fig. 4), where valley widening has occurred at the expense of deepening. Histogram from Haramosh included for comparison with a. c, Distribution of slope angles for largely intact and partially dissected plateaux. The central Deosai plateau (area 7,

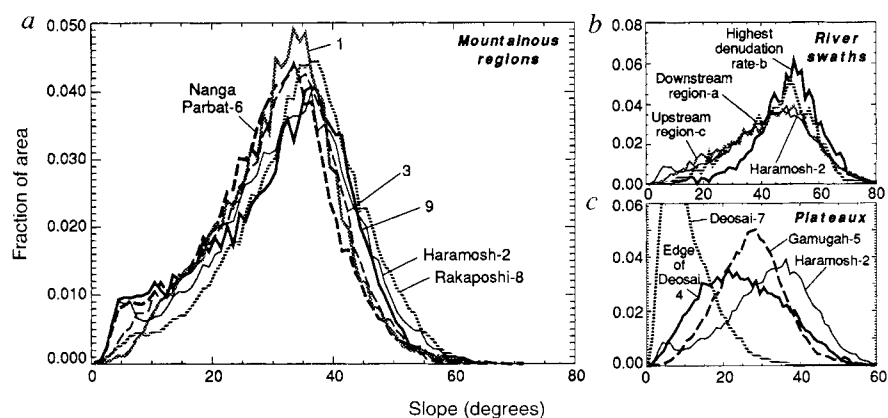


Fig. 4) is dominated by low slopes, whereas those areas including parts of dissected plateaux (areas 4 and 5) show intermediate slope distributions compared to nearby mountainous zones. Histogram from Haramosh included for comparison with a.

sediment means that hillslope angles are not diminished by valley aggradation. Because most of the bedrock is extensively fractured, it lacks cohesion at the scale of large hillslopes. Thus, we interpret the typically high mean slope angles ( $32^\circ$  using a  $270 \times 270$ -m window) as reflecting an average effective angle of internal friction which controls hillslope stability in this fractured bedrock.

Regardless of the controlling process of denudation, the similarity of slope distributions and the comparable mean altitudes (4,200–4,600 m) based on hypsometric (area versus altitude) analysis of all mountainous regions suggest homogeneous topographic characteristics, largely independent of denudation variations, in this landscape. The association of high river gradients with high rock-uplift rates implies a dynamic equilibrium in which bedrock uplift and incision are balanced through adjustments in stream power. Rates of denudation are sufficiently high that, even in areas with denudation rates of  $1\text{--}2 \text{ mm yr}^{-1}$ , the processes that typically affect soil-mantled slopes<sup>34</sup> (regolith production, creep and slope wash) are volumetrically unimportant. Instead, a threshold incision rate (perhaps  $0.5 \text{ mm yr}^{-1}$ ) has been crossed, whereby rapid mass transfer through landsliding is the primary means by which the hillslopes adjust to changes in boundary conditions resulting from river incision at their toes. In this case, mean slope angles will be set by rock strength, and mean relief will be

controlled by the spacing of large rivers<sup>26</sup>. Given that the spacing of large rivers is similar among different subregions, and that denudation calculations indicate at least several kilometres of bedrock erosion during the past 1 Myr throughout most of the area, attainment of dynamic equilibrium and uniform slope and altitudinal characteristics is perhaps not surprising. Herein, we have actually documented such an equilibrium for a very large-scale, high-relief and rapidly deforming part of the largest mountain range in the world.

A conceptual model for mountainous regions of variable but persistently high rates of bedrock uplift emerges from this analysis. Here, large rivers can attain equilibrium profiles wherein local fluvial gradients are adjusted until incision rates driven by stream power balance bedrock uplift rates. Large rivers become bedrock saws, fixed with respect to the geoid. Between rivers, bedrock rises towards denuding hillslopes whose rate of mass transfer to valley bottoms through landsliding is proportional to the incision rates of the great rivers. Because the threshold for landsliding is determined by fractured rock strength, the highest topography will develop either where rocks are less fractured or where large rivers are most widely spaced. To test or quantify this model, more data are needed on mass-flux rates through bedrock uplift and landsliding, on rates of river incision and glacial erosion, and on topographic variability within deforming regions. □

Received 18 August; accepted 27 December 1995.

- Molnar, P. & England, P. *Nature* **346**, 29–34 (1990).
- England, P. & Molnar, P. *Geology* **18**, 1173–1177 (1990).
- Misch, P. *Festschrift zum 60. Geburtstag von Hans Stille*, 259–276 (Enke, Stuttgart, 1936).
- Butler, R. W. H. & Prior, D. J. *Nature* **333**, 247–250 (1988).
- Seeber, L. & Gornitz, V. *Tectonophysics* **92**, 335–367 (1983).
- Zeitler, P. K., Johnson, N. M., Naeser, C. W. & Tahirkehl, R. A. *Nature* **298**, 255–257 (1982).
- Zeitler, P. K. *Tectonics* **4**, 127–151 (1985).
- Treloar, P. J. et al. *Tectonics* **8**, 881–909 (1989).
- Zeitler, P. K., Chamberlain, C. P. & Smith, H. *Geology* **21**, 247–250 (1993).
- Winslow, D. M., Zeitler, P. K., Chamberlain, C. P. & Hollister, L. S. *Geology* **22**, 1075–1078 (1994).
- George, M., Reddy, S. & Harris, N. *Tectonics* **14**, 237–252 (1995).
- Sorkhabi, R. B. *Nucl. Tracks Radiat. Meas.* **21**, 535–542 (1993).
- Nishizumi, K. et al. *J. geophys. Res.* **94**, 17907–17915 (1989).
- Kohl, C. P. & Nishizumi, K. *Geochim. cosmochim. Acta* **56**, 3583–3587 (1992).
- Lai, D. *Earth planet. Sci. Lett.* **104**, 424–439 (1991).
- Nishizumi, K. et al. *Earth Surf. Proc. Landforms* **18**, 407–425 (1993).
- Naeser, N. D., Naeser, C. W. & McCulloh, T. H. In *Thermal History of Sedimentary Basins: Methods and Case Studies* (eds Naeser, N. D. & McCulloh, T. H.) 157–180 (Springer, New York, 1989).
- Harrison, T. M. *Nucl. Tracks* **10**, 329–333 (1985).
- Naeser, C. W. *Nucl. Tracks* **5**, 248–250 (1981).
- Stüwe, K., White, L. & Brown, R. *Earth planet. Sci. Lett.* **124**, 63–74 (1994).
- Seidl, M. A. & Dietrich, W. E. *Catena Suppl.* **23**, 101–124 (1992).
- Howard, A. D., Dietrich, W. E. & Seidl, M. A. *J. geophys. Res.* **99**, 13971–13986 (1994).
- Rosenbloom, N. A. & Anderson, R. S. *J. Geophys. Res.* **99**, 14013–14029 (1994).

- Cronin, V. S. In *Tectonics of the Western Himalayas* (eds Malinconico, L. L. Jr & Lillie, R. J.) 181–202 (Spec. Pap. 232, Geol. Soc. Am., Boulder, Colorado, 1989).
- Cronin, V. S. & Johnson, G. D. In *Himalaya to the Sea Geology, Geomorphology and the Quaternary* (ed. Schroder, J. F. Jr) 91–107 (Routledge, London, 1993).
- Anderson, R. S., Dick, G. S. & Densmore, A. *Geol. Soc. Am. Abstr. Prog.* **26**, 238–239 (1994).
- Sorkhabi, R. B. et al. *Indian Acad. Sci. (Earth planet. Sci.)* **103**, 83–97 (1994).
- Penck, W. *Morphological Analysis of Landforms* (Macmillan, London, 1953).
- Hanks, T. C., Bucknam, R. C., Lajoie, K. R. & Wallace, R. E. *J. geophys. Res.* **89**, 5771–5790 (1994).
- Dietrich, W. E., Wilson, C. J., Montgomery, D. R. & McKean, J. J. *Geol.* **101**, 161–180 (1993).
- Fielding, E. J., Isacks, B. L., Barazangi, M. & Duncan, C. *Geology* **22**, 163–167 (1994).
- Carson, M. A. & Kirby, M. J. *Hillslope Form and Process* (Cambridge Univ. Press, Cambridge, 1972).
- Schroder, J. F. Jr, Khan, M. S., Lawrence, R. D., Madin, I. P. & Higgins, S. M. In *Tectonics of the Western Himalayas* (eds Malinconico, L. L. Jr & Lillie, R. J.) 275–294 (Spec. Pap. 232, Geol. Soc. Am., Boulder, Colorado, 1989).
- Dietrich, W. E., Wilson, C. J., Montgomery, D. R., McKean, J. & Bauer, R. *Geology* **20**, 675–679 (1992).

ACKNOWLEDGEMENTS. Support for this research was provided by the National Science Foundation, National Geographic Research and the National Aeronautics and Space Administration. We thank J. Vergés and J. Khan for assistance with field work and logistics, B. Isacks for discussions and contributions to the DEM analysis, K. Hodges, A. Howard and M. Seidl for reviews, M. Caffee, R. Finkel and J. Southon for assistance with the cosmogenic nuclide analyses at Lawrence Livermore National Laboratory, Earth Resource Mapping for an educational grant and Sun Microsystems for donating a graphics tower.



# The structure of the *Escherichia coli* EF-Tu·EF-Ts complex at 2.5 Å resolution

Takemasa Kawashima\*, Carmen Berthet-Colominas\*, Michael Wulff†, Stephen Cusack\* & Reuben Leberman\*‡

\* European Molecular Biology Laboratory, Grenoble Outstation, B.P. 156, 38042 Grenoble Cedex, France

† European Synchrotron Radiation Facility, B.P. 220, 38043 Grenoble Cedex, France

**The crystal structure of the EF-Tu·EF-Ts complex from *Escherichia coli* has been determined to a resolution of 2.5 Å. The complex contains two subunits of each of the elongation factors. The two EF-Ts molecules form a tight dimer, but there is little contact between the two EF-Tu molecules. The interaction of EF-Ts with EF-Tu results principally in the disruption of the Mg<sup>2+</sup> ion binding site, thereby reducing the affinity of EF-Tu for guanine nucleotides.**

BACTERIAL polypeptide elongation factor, EF-Tu, is the archetypal guanine-nucleotide-binding protein<sup>1</sup> (G protein). In bacterial protein biosynthesis, it transports aminocylated elongator transfer RNAs to the messenger RNA:ribosome complex. This essential role led to it being studied intensively many years before the large extended family of GTPases and their importance in the regulation of a wide range of metabolic processes were discovered<sup>2</sup>. All of these proteins have functions for which analogous reaction schemes can be described. They involve the hydrolysis of bound GTP to GDP during the regulatory process in which the protein-GDP complex must be recycled to the protein-GTP complex. This recycling of inactive GDP complexes to the active GTP complexes is catalysed by guanine-nucleotide exchange factors, which prevent the G protein from being locked in the inactive state.

In the case of EF-Tu, the guanine-nucleotide exchange factor is a soluble protein factor called EF-Ts<sup>3</sup>. In the classical reaction scheme<sup>4</sup> for the elongation step in bacterial polypeptide synthesis, EF-Ts acts as a catalyst in the displacement of the GDP from the EF-Tu·GDP complex and allows the binding of GTP so that the ternary complex EF-Tu·GTP·aminoacyl-tRNA can be formed. The catalytic function of EF-Ts is consistent with the relative amounts the two factors present in the bacterial cell, and is required because of the hundred-fold higher affinity of EF-Tu for GDP than for GTP<sup>5</sup>. Despite the high affinity between the two elongation factors, in *E. coli* the EF-Tu·EF-Ts complex appears to have a transient existence because it is not observed on direct fractionation of cell extracts<sup>6</sup>. It may, however, be more stable in bacteriophage-infected cells as a component of the RNA-dependent RNA polymerases of bacteriophages, typified by Q $\beta$ -replicase<sup>7</sup>.

Analysis of the crystal structure of the trypsin-modified EF-Tu·GDP<sup>8</sup> complex of *E. coli* shows that it is a rather flat triangular molecule with an unusually large hole<sup>9</sup>, and consists of three structural domains. Domain 1, the amino-terminal domain, is responsible for the binding on the guanine nucleotide, and is structurally similar to that of the Ha-Ras oncogene product, p21 (ref. 10). Domain 2 has the structure of a  $\beta$ -barrel with seven antiparallel strands, and is connected to domain 1 through a single flexible peptide 16 Å long. Domain 3 also has a  $\beta$ -barrel structure, with six antiparallel strands, is connected to domain 2 by a short peptide chain, and appears to be tightly associated with domain 1 (ref. 8). The crystal structures of EF-Tu from *Thermus aquaticus* and *T. thermophilus* complex with the GTP analogue GDPNP<sup>11,12</sup> show that there are striking conformational differ-

ences between the protein in the GTP form compared to that in the GDP form. In the GTP form, domain 2 packs tightly against domain 1 so the large central hole is not present.

The relative molecular mass ( $M_r$ ) of EF-Tu is 43.2K and EF-Ts 30.3K<sup>14</sup>. The EF-Tu·EF-Ts complex was crystallized from polyethylene glycol solutions<sup>15</sup> and a preliminary X-ray diffraction analysis was undertaken many years ago<sup>16</sup>. Here we describe the three-dimensional structure determination of the complex, using crystals obtained under similar conditions except that the crystals were converted and maintained in a high-resolution form<sup>17,18</sup>, rapid freezing techniques were used, and we had access to the high-intensity X-ray sources at the European Synchrotron Radiation Facility (ESRF). The structure suggests a mechanism for the guanine-nucleotide release/exchange for EF-Tu, which may be similar for other guanine-nucleotide-binding proteins.

## Data collection

Purification of the EF-Tu·EF-Ts complex, crystallization and preparation of the crystals for data collection are described in detail elsewhere<sup>18</sup>. Crystals were grown by the sitting-drop technique from a solution with a final composition of 23% PEG 6000, 20% PEG 400, 100 mM KNO<sub>3</sub>, 64.4 mM Tris-acetate at pH 7.7 at 19–20 °C. These crystals belonged to the space group  $P2_12_12_1$  with cell parameters  $a = 74.5$  Å,  $b = 109.6$  Å and  $c = 198.8$  Å. Two heavy-atom derivatives were obtained by soaking crystals in the same solution with 750  $\mu$ M trimethyl lead acetate for 48 h, or 10  $\mu$ M ethyl mercury chloride for 20 h. All diffraction measurements were performed at  $\sim -163$  °C on flash-frozen crystals of approximate dimensions  $400 \times 300 \times 200$   $\mu$ m<sup>3</sup> which diffracted beyond 2.4 Å. The cell parameters were reduced to  $a = 73.5$  Å,  $b = 108.5$  Å and  $c = 194.5$  Å at low temperature. Native, lead and mercury derivative data sets were collected using one crystal per data set, at ESRF on beam-line 3 (ID9), with a 72-pole (46-mm period) undulator source and a fundamental at 2.37 keV and 30-cm Mar-Research Image plate system as detector.

## Model quality

The data used for the structure determination are summarized in Table 1, with an example of the electron-density map shown in Fig. 1. Using all data between 2.5 Å and 8 Å resolution, the current model has a crystallographic  $R$ -factor of 17.5% and a free  $R$ -factor of 28.3%, with a root mean square (r.m.s.) deviation from ideality of 0.007 Å for bond lengths and 1.42 ° for bond angles, with 1,183 water molecules. The average  $B$ -factor is 37.3 Å<sup>2</sup>, where 36.6 Å<sup>2</sup> accounts for the complexes and 43.5 Å<sup>2</sup> for the water molecules. The Ramachandran plot displays two outliers for one of the EF-Ts molecules (monomer a), Gln 281 and Glu 268, the latter being

‡ To whom correspondence should be addressed.

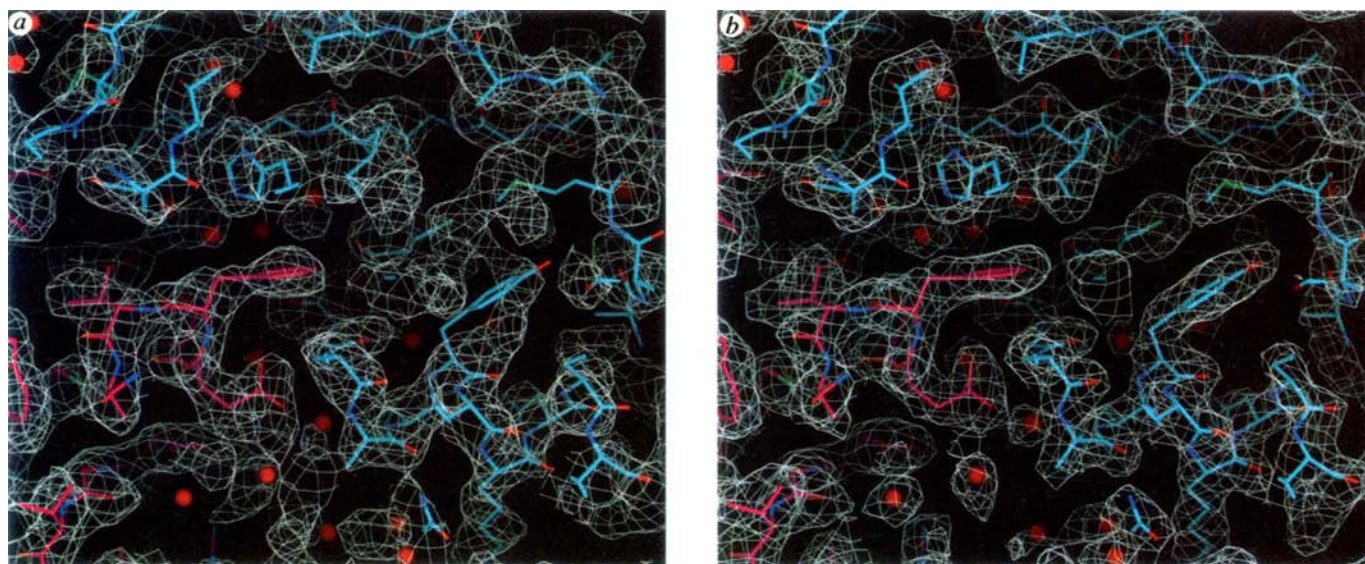


FIG. 1 Examples of electron-density map, superimposed with the current model. The peptide chain is shown in cyan for EF-Tu and magenta for EF-Ts; red spheres represent water molecules. The maps are contoured at  $1\sigma$ , in a region where the conserved Asp 80 and Phe 81 of EF-Ts (in the centre) are interacting with helix B of EF-Tu (on the right). *a*, The initial

n.c.s.-averaged map. Note that the density is already reasonable, except around the tyrosine and the methionine residues shown on the right. *b*, The final  $2f_o - f_c$  map. This figure was prepared using O<sup>26</sup>.

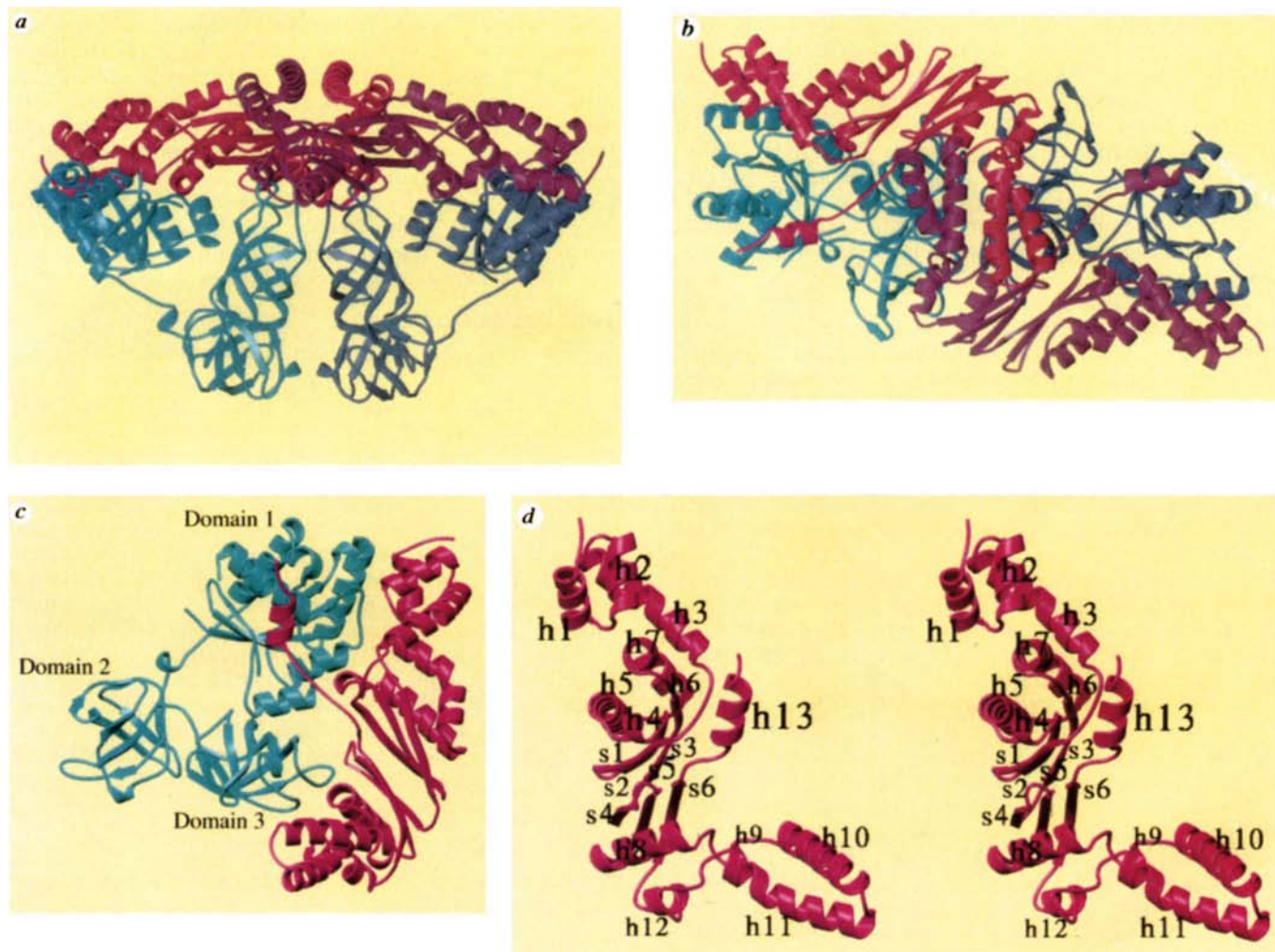


FIG. 2 Overall structure of the EF-Tu-EF-Ts complex. *a*, The dimer with the non-crystallographic pseudo-two-fold axis oriented vertically. The EF-Tu molecules are represented in cyan and slate blue, the EF-Ts in magenta and purple. The nucleotide binding domain made of alternate  $\alpha$ -helices and  $\beta$ -strands as well as the tip of the  $\beta$ -barrel domain 3 are in contact with EF-Ts. Only minor contacts can be observed between EF-Tu molecules. *b*, View of the dimer down the n.c.s. axis, that is, after a  $90^\circ$  rotation about the horizontal axis from *a*. The two antiparallel helices in the centre display the largest interface between monomers, and are

responsible for dimerization of the EF-Tu-EF-Ts complex. *c*, View of a monomer of the complex showing that the amino-terminal half of EF-Ts interacts with the nucleotide binding domain 1 of EF-Tu, and the other half binds to domain 3, with the exception of the helical hairpin (at the bottom of the picture) which is involved in dimerization, and the C-terminal module which packs against domain 1 of EF-Tu. *d*, Stereo view of EF-Ts viewed from the EF-Tu perspective. The secondary structure elements are indicated. This figure was prepared with MOLSCRIPT<sup>28</sup> and RASTER3D<sup>29</sup>.

part of a peptide linking the helical appendix packed against domain 1 of EF-Tu and the core of EF-Ts. Residues Ser 221, Ile 247 and Arg 333 of EF-Tu also lie outside the allowed areas. They are all involved in turns from one  $\beta$ -strand to the next, in the vicinity of, or directly involved in, crystal contacts. Also, Gln 97 of monomer b has a poor geometry owing to alternative conformations and confused electron density in the region Ala 95 to Gln 97 of EF-Tu. The last three amino acid residues of EF-Ts are weakly defined, as are loops 219–224, 255–266 and 281–285, in domain 2 of EF-Tu, especially for monomer b. In EF-Tu, we have not modelled residues 42–63, including the effector loop (some scattered densities are visible), or the first eight residues (for which no density can be seen).

## Overall structure

The asymmetric unit found in the crystal structure is composed of two molecules of EF-Tu and two of EF-Ts. They are organized as

a dimer of the EF-Tu·EF-Ts complex, with the monomers related by a non-crystallographic pseudo-two-fold axis (Fig. 2a). The dimension along this axis is 80 Å, the width is 140 Å and the depth is 70 Å. The major contact between the two monomers is by an antiparallel helical hairpin protruding from EF-Ts which interacts with its non-crystallographic symmetry (n.c.s.)-related counterpart in the other EF-Ts subunit (Fig. 2b). There are smaller contacts between domains 3 of the EF-Tu subunits. These observations suggest that EF-Ts is responsible for dimerization, and that the quaternary structure is best described by [EF-Ts]<sub>2</sub>·2EF-Tu. The conformation of EF-Tu bound to EF-Ts is similar to that in the EF-Tu·GDP complex<sup>8</sup>. EF-Ts is an elongated molecule and binds principally to one EF-Tu, spanning domain 1 and the tip of domain 3 (Fig. 2c, d). There are also minor contacts between the tip of domain 3 of EF-Tu belonging to one monomer and the EF-Ts from the other. In the following, the molecules will be labelled aTu (aEF-Tu) and aTs (aEF-Ts) for those forming the

TABLE 1 Summary of crystallographic data

| (a) Data collection statistics |                     |                        |                      |                        |
|--------------------------------|---------------------|------------------------|----------------------|------------------------|
|                                | Native              |                        | Mercury              | Lead                   |
| Resolution limits (Å)          | 43.67–2.48          |                        | 43.66–2.47           | 42.51–2.48             |
| Cell parameters (Å)            |                     |                        |                      |                        |
| <i>a</i>                       | 73.53               |                        | 73.49                | 73.37                  |
| <i>b</i>                       | 108.54              |                        | 108.53               | 108.26                 |
| <i>c</i>                       | 194.55              |                        | 194.75               | 194.63                 |
| Total reflections              | 210,876             |                        | 185,903              | 138,537                |
| Independent reflections        | 54,834              |                        | 53,732               | 45,277                 |
| Redundancy                     | 3.8 (3.2)*          |                        | 3.5 (2.6)*           | 3.1 (2.4)*             |
| Completeness                   | 98.2 (80.3)*        |                        | 95.3 (74.2)*         | 80.7 (60.9)*           |
| <i>I</i> / <i>SigI</i>         | 4.9*                |                        | 3.5*                 | 4.3*                   |
| <i>R</i> <sub>sym</sub>        | 4.3 (15.1)*         |                        | 4.4 (20.5)*          | 3.3 (15.5)*            |
| (b) MIR phase statistics       |                     |                        |                      |                        |
|                                | Mercury             |                        |                      | Lead                   |
|                                | Relative occupancy  | Binding site           | Relative occupancy   | Binding site           |
| Site 1                         | 0.17                | uaCys 81<br>uaMet 98   | 0.29                 | saGlu 193<br>sbGlu 193 |
| Site 2                         | 0.15                | ubCys 81<br>ubMet 98   | 0.25                 | ubGlu 348              |
| Site 3                         | 0.12                | uaCys 137<br>uaLeu 146 | 0.20                 | uaGlu 348              |
| Site 4                         |                     |                        | 0.15                 | saGlu 13               |
|                                | Centric reflection  | Acentric reflection    | Centric reflection   | Acentric reflection    |
| Phasing powers                 | 0.53                | 0.72                   | 0.63                 | 0.83                   |
| Cullis <i>R</i> -factor        | 0.87                | 0.93                   | 0.82                 | 0.89                   |
|                                | Centric reflections |                        | Acentric reflections | Mean                   |
| Figure of merit                | 0.49                |                        | 0.31                 | 0.33                   |

The MOSFLM package<sup>25</sup> was used for integration; the CCP4<sup>19</sup> suite was used for data reduction, scaling, phase calculation and density modification; the O/RAVE package<sup>26,27</sup> was used for model building and for map, mask and bone manipulation; and the refinement was carried out using X-PLOR<sup>27</sup>. Heavy-atom sites were found with the help of the programme RSPS<sup>19</sup> and cross- and double-difference Fourier. All sites were refined and multiple isomorphous replacement (MIR) phases calculated using MLPHARE<sup>18</sup> followed by phase improvement with the solvent flattening option of DM<sup>19</sup>. Careful inspection of the skeletonized map allowed the location of a strand–turn–helix near one lead site (Pb-3). This structural element could not be matched with any features of EF-Tu, and was therefore attributed to EF-Ts. A similar element was found near another lead site (Pb-2). The main lead site (Pb-1) was at a position equidistant from Pb-2 and Pb-3, suggesting that a non-crystallographic two-fold axis passed through Pb-1. This n.c.s. operator combined with a mask based on selected bones was input into the molecular averaging option of DM. In the resulting maps, more secondary structure elements could be identified, which in turn were included in the mask. This procedure was repeated until much of EF-Ts and then domain 3 of EF-Tu became visible, allowing the placing of the EF-Tu·GDP model<sup>8</sup> in the map. The model building of EF-Ts and the rebuilding of EF-Tu involved extensive use of the Real Space Refinement implemented in O<sup>26</sup>. An ambiguity as to whether some helical density packed against domain 1 of EF-Tu was part of the effector loop of EF-Tu or the C-terminal extension of EF-Ts was resolved by trying to refine both possibilities and comparing the results to a real space averaged map for the region calculated with no model. This showed that the density corresponded to EF-Ts, which was confirmed in subsequent refinement steps. The final model was obtained by the following refinement protocol: starting from a model with strict non-crystallographic symmetry, n.c.s. restraints were used in conjunction with manual corrections to reduce the free R-factor in a series of positional and B-factor refinement cycles. At a stage where the R-factor was around 21% (free-R ~ 31%), water molecules were gradually added, 40–90 at a time. No more simulated annealing was then performed, and waters with a B-factor > 75 Å<sup>2</sup> were rejected, as were those with a real space correlation below 50%. In the latest refinement steps, the n.c.s. restraints were no longer used.

\* Between 2.55 Å and highest resolution.



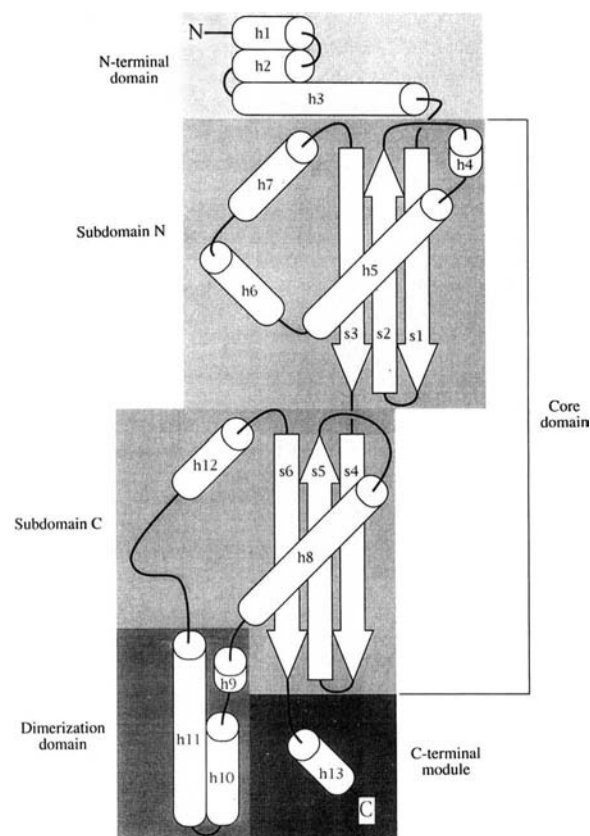
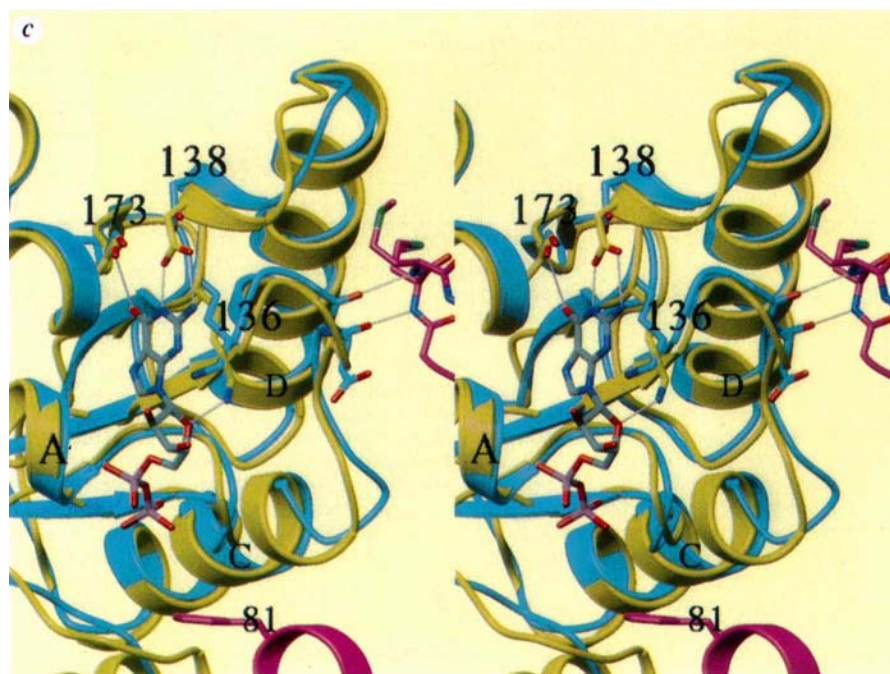
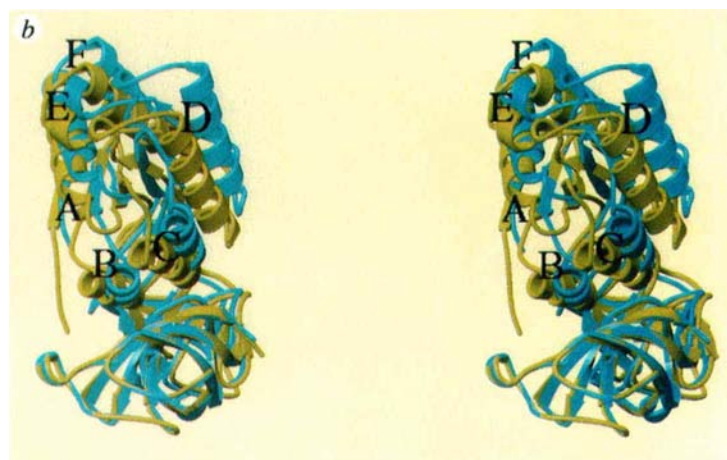
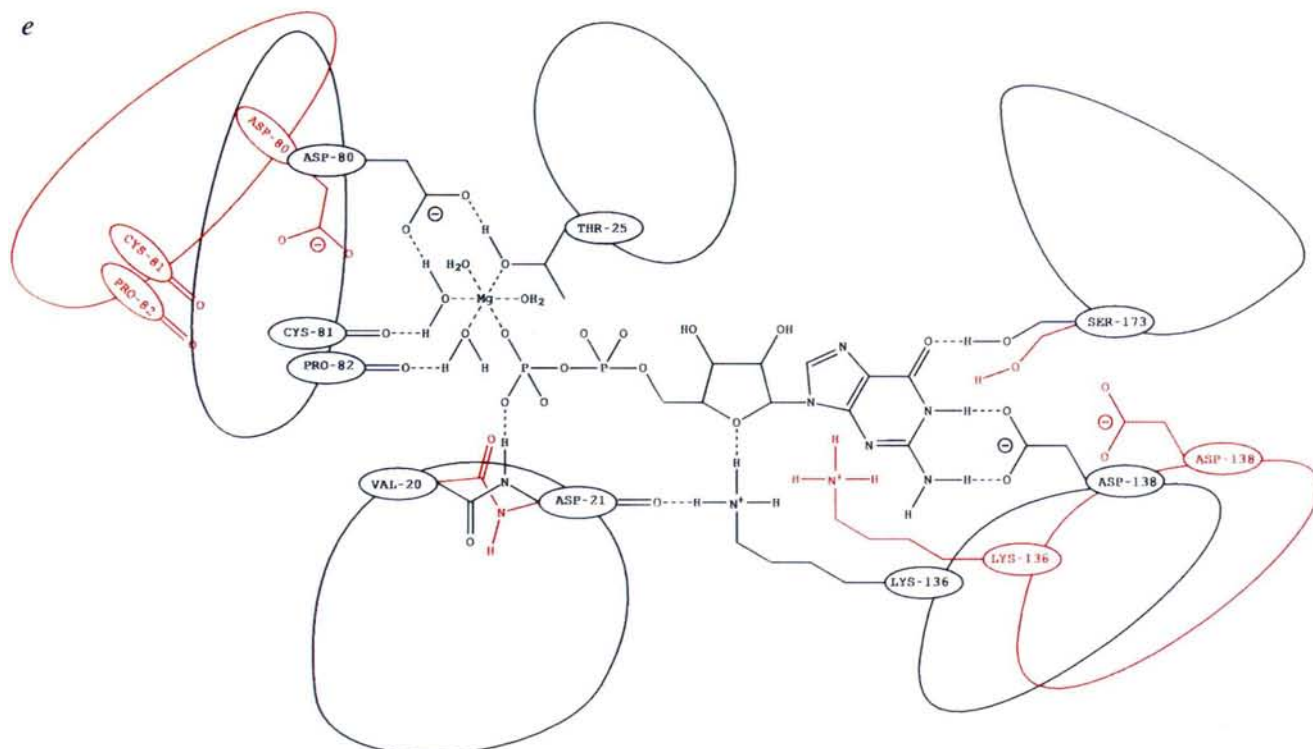
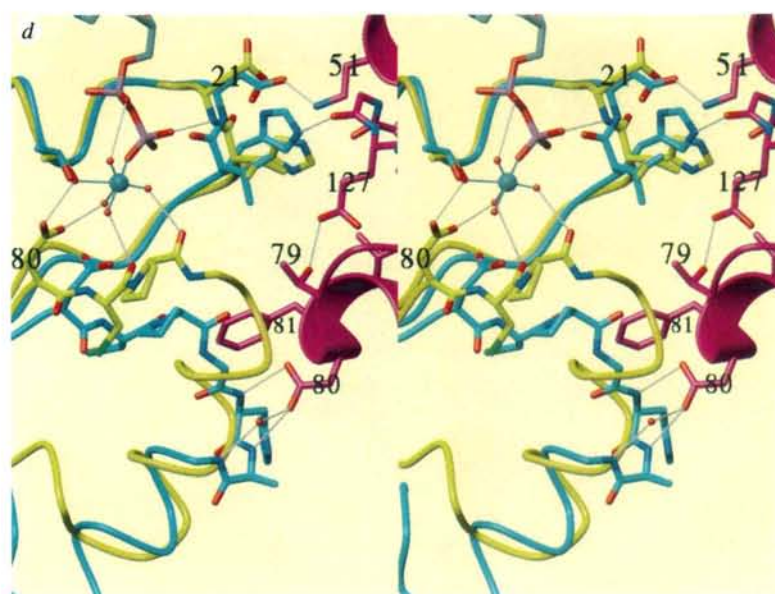
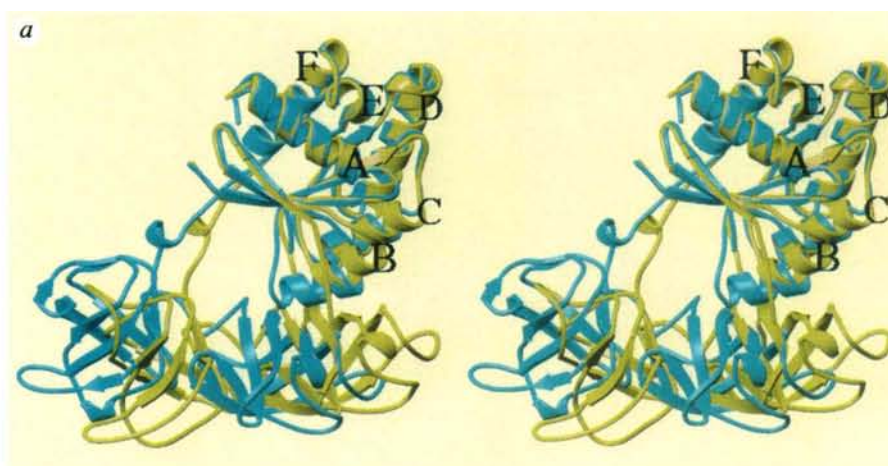
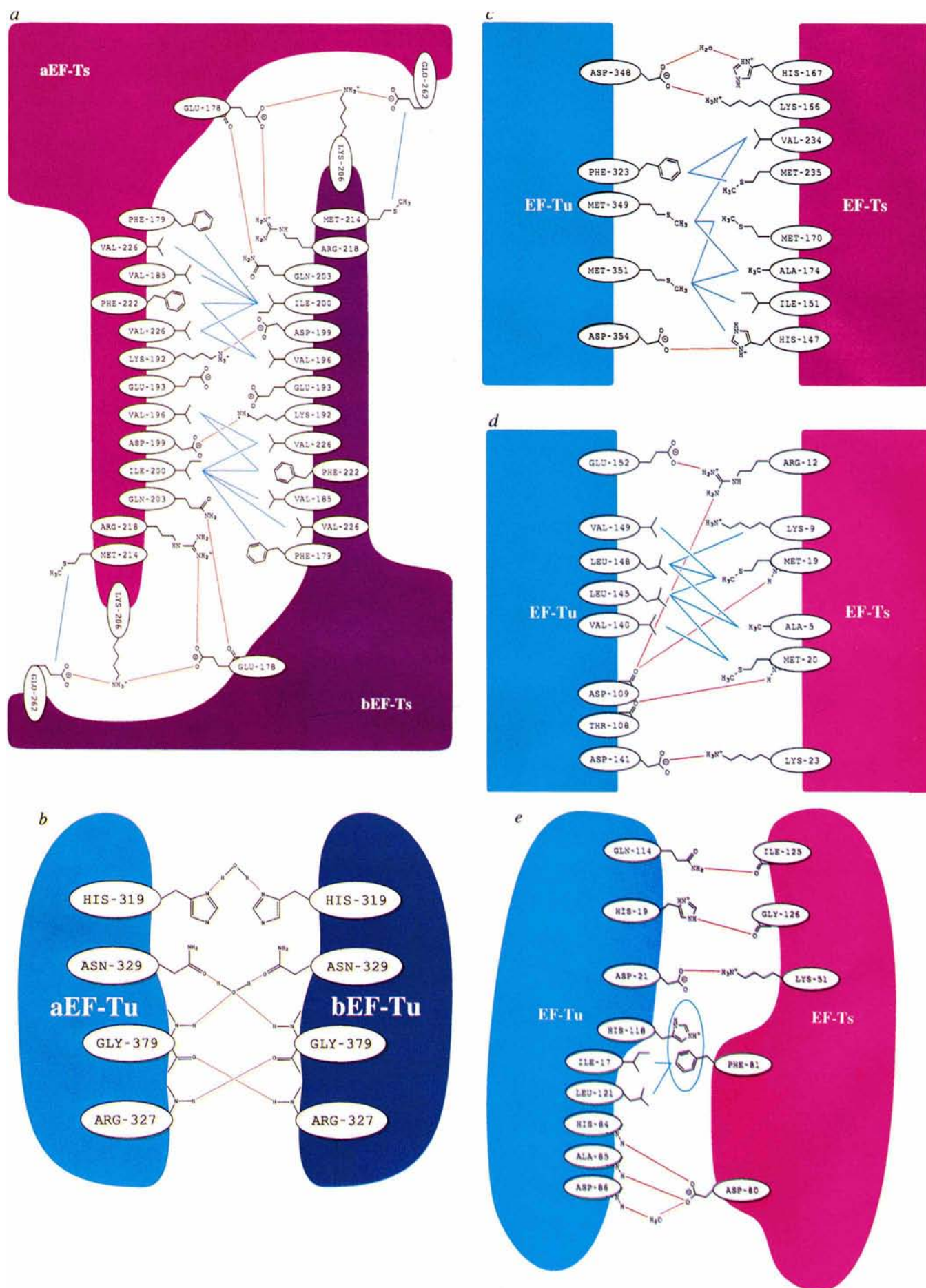


FIG. 4 Structural rearrangement between the GDP and EF-Ts bound forms of EF-Tu. The EF-Tu in the GDP complex is shown in yellow, and in the EF-Ts complex in cyan. The helices of EF-Tu are labelled. *a*, Global movements relative to domain 1. The two structures have been superimposed by their respective helices A and F with the least-square explicit function of O. Almost every  $\beta$ -strand of domain 1 is then superimposed, whereas helix B has shifted as a rigid body by one helical turn. Domains 2 and 3 have also moved as a rigid body by a rotation of about  $18^\circ$ . All discussions about the nucleotide binding site are based on this three-dimensional alignment. *b*, Global movement relative to domain 3. The orientation is  $85^\circ$  apart around the vertical axis with respect to the previous figure; the two structures were superimposed by their respective domain 3. The view is almost down the axis of the B helix in the EF-Ts bound in addition to EF-Tu. The movement relative to domain 3 for the B helix is a torsion towards the D side, whereas the rest of the domain 1, the twist, has lifted from domain 3. *c*, Changes in the vicinity of the guanine base. The GDP is represented with a sea-green backbone. Oxygens are in red, nitrogens blue, phosphorus purple, and EF-Ts backbone magenta. The hydrogen bonds are symbolized by thin grey lines. The interaction of the loop upstream of helix C is represented in the top right of the picture. The intrusion of sPhe 81 is shown at the bottom. The strand in the bottom left is the switch loop, followed by helix B. *d*, Changes in the vicinity of the magnesium binding site. The orientation is slightly different from *c*. In addition, water molecules are represented by small red spheres, and steel grey is used for the magnesium ion and the bonds symbolizing its coordination. The C and D helices have been omitted for clarity. *e*, Schematic representation summarizing the modifications of the nucleotide binding site. The changes are indicated in red. *a*–*d* were prepared using MOLSCRIPT<sup>28</sup> and RAS-TER3D<sup>29</sup>.











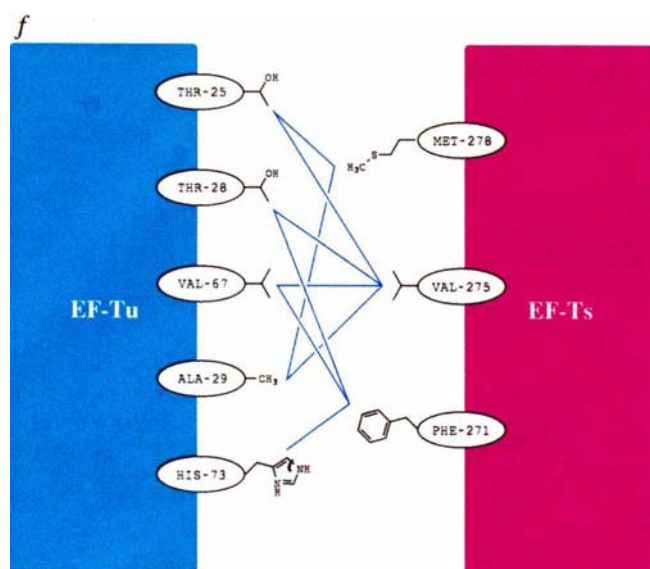


FIG. 5 Schematic representation of the intermolecular contacts in the complex. The hydrophobic interactions are indicated by blue lines, polar and electrostatic interactions are in red. *a*, Ts–Ts interactions; *b*, Tu–Tu interactions; *c*, interactions between EF-Tu domain 3 and EF-Ts subdomain C; *d*, interactions involving the N-terminal domain of EF-Ts with the nucleotide binding domain 1 or EF-Tu; and *e*, interactions of the subdomain N with the nucleotide binding domain. The blue circle indicates an interaction where the plane of the imidazole ring is stacked perpendicular to the plane of the phenyl ring. Note that sLys 51, although involved in the interaction with the phosphate binding loop, is not strictly part of subdomain N. *f*, the C-terminal module.

first complex, and bTu (bEF-Tu) and bTs (bEF-Ts) for the second. The residues will be referred to by prefixing their names and numbers with 'u' if they belong to EF-Tu and 's' if they belong to EF-Ts (for example, uMet 349 and sMet 235), with the appropriate modifier, a or b (for example, saGlu 236 and ubArg 318). We will also refer to the 'D' side, the side of the triangle formed by the three domains of EF-Tu to which the D helix belongs, as opposed to the side which includes the nucleotide binding site.

### EF-Ts moiety

EF-Ts is composed of four structural modules: the N-terminal domain (residues 1–54), the core (55–179 and 229–263), the dimerization insertion (180–228) and the C-terminal module (264–282). The core domain can be further divided into N-terminal and C-terminal subdomains. Subdomain N (residues 55–140) contains a secondary structure motif which is repeated in subdomain C (residues 141–179 and 229–263), so that the core part of EF-Ts shows an internal pseudo-symmetry (Fig. 3). The motif has a three-stranded antiparallel  $\beta$ -sheet, and one  $\alpha$ -helix: s1, s2, s3 and h5 in subdomain N, corresponding to s4, s5, s6 and h8 in subdomain C, respectively. These secondary structure elements are organized as helix–sheet–sheet–helix where the  $\beta$ -sheet from subdomain N faces the  $\beta$ -sheet from subdomain C, with a cross-over angle of approximately  $20^\circ$ ; each  $\alpha$ -helix lies on its respective side in a plane parallel to that of the proximal sheet, with an angle of  $\sim 45^\circ$  between the axis of the helix and the sheet. The r.m.s. deviation of the backbone after superposition of these elements within one EF-Ts molecule is 1.5 Å for 37  $\alpha$ -carbons. Other helices are present in more distal positions (h7 and h12), roughly following the same internal symmetry (Figs 2 and 3). A protruding antiparallel coiled-coil (h10 and h11), showing no contacts other than with its n.c.s.-related partner in the other EF-Ts monomer, is inserted in subdomain C. The dimerization domain, residues 180–228, consist of this coiled-coil and the preceding helix h9. The EF-

Ts dimer interface involves three charged residues from each monomer (Lys 192, Glu 193 and Asp 199), as well as hydrophobic contacts (Phe 179, Val 226, Val 185, Phe 222, Val 189, Val 196 and Ile 200). The tip of the helical hairpin of one EF-Ts is in contact with the core of the other EF-Ts. The centre of symmetry of all these interactions is Glu 193. The N-terminal domain is composed of two  $\alpha$ -helices (h1 and h2) and a third long, bent helix (h3) lying over the EF-Tu side of subdomain N, connected to the centre of the core domain. The C-terminal module of EF-Ts consists of a linker peptide followed by an  $\alpha$ -helix (h13) which binds to domain 1 of EF-Tu.

### EF-Tu moiety

Superposing the EF-Tu molecules of the EF-Ts and GDP complexes by domain 1 shows at the quaternary level that domains 2 and 3 swing away from domain 1 by approximately  $18^\circ$ , with an average displacement of about 10 Å (Fig. 4*a*). If the superposition is by domain 3, domain 2 shows very little movement with respect to domain 3. At the tertiary level, the most significant change is that of the module between residues 80–97 of domain 1, which swings, essentially as a rigid body hinged at its extremities (Val 79 and Met 98), such that the helix B is displaced by one helical turn in the C-terminal direction (Fig. 4*a*). This helix is known to be tightly associated to domain 3 (refs 8, 11, 12), and the relative movement with respect to the latter is a twist of  $25$ – $30^\circ$  towards the D-side, in contrast to the rest of domain 1 which is swung in the same direction by  $25^\circ$  and lifted from domain 3 (Fig. 4*b*). While residues 80–93 move principally as a rigid body, the strand encompassing residues 94–96 becomes disordered, the electron density in this region showing two conformations  $180^\circ$  apart around Ala 96 and Gln 97.

The helical turn of the peptide that links domain 1 to domain 2 is shifted in the primary structure, spanning residues 205–207 in the complex with EF-Ts as compared to residues 203–205 in the complex with GDP. This brings it closer to domain 2, as a result of the reorganization between domains. The loops 257–268 and 280–285 of domain 2, involved in crystal packing in the EF-Tu·EF-Ts structure, also adopt alternate backbone traces. The EF-Tu monomers have some contacts between each other, but involving residues of domain 3 only (Fig. 5*b*) where the backbone NH of uaArg 327 is connected by hydrogen bonds to the backbone oxygen of ubGly 379, and the backbone NH of the latter and the amide oxygen of ubAsn 329 have hydrogen bonds, by means of a water molecule, to their non-crystallographically related counterparts (all of these interactions are symmetrical). A similar water mediation is seen between uaHis 319 and ubHis 319.

### EF-Tu EF-Ts contacts

Domain 3 of EF-Tu is bound to subdomain C of the core of EF-Ts, whereas domain 1 of EF-Tu interacts with subdomain N, the N-terminal domain and the C-terminal module of EF-Ts. The decrease in accessible surface, calculated using SURFACE from CCP4 (ref. 19), between the molecules in the absence and in the presence of different partners in the complex is  $2,224 \text{ Å}^2$  for EF-Tu, of which  $1,700 \text{ Å}^2$  is due to the interaction with one EF-Ts,  $463 \text{ Å}^2$  to the contact with the other EF-Tu, and  $142 \text{ Å}^2$  to the cross-contact with the second EF-Ts; the decrease is  $3,092 \text{ Å}^2$  for EF-Ts, of which,  $1,886 \text{ Å}^2$  is due to the contact with EF-Tu,  $1,070 \text{ Å}^2$  from dimerization with EF-Ts, and  $137 \text{ Å}^2$  is due to the cross-interaction with EF-Tu. The intermolecular interactions are schematically represented in Fig. 5.

The most significant feature of the EF-Tu·EF-Ts contact is situated in subdomain N, where sPhe 81, a residue of the conserved TDFV sequence motif found in all EF-Ts, intrudes between the side chains of uHis 84 and uHis 118, belonging to helices B and C, respectively, of the EF-Tu molecule. The shifted helix B appears to be finely oriented by the conserved sAsp 80, capping in an intermolecular fashion the amino terminus of the helix at the peptide nitrogens of uHis 84, uAla 85 and uAsp 86, the last interaction being mediated by a water molecule. The peptide

oxygens of sIle 125 and sGly 126 (residues of a conserved  $\Phi$ GEN $\Phi$  motif) are linked by hydrogen bonds to the amide nitrogen of uGln 114 (part of helix C) and the  $\epsilon$ -nitrogen of uHis 19 (part of the phosphate binding loop), respectively, maintaining the displacements provoked by the sPhe 81 intrusion. In addition, sLys 51, held in place by the amide oxygen of sGln 54, interacts with the phosphate binding loop by means of the carboxyl group of uAsp 21 (Figs 4d and 5e).

### Nucleotide release/exchange

It has been known for many years from solution studies<sup>20</sup> that removal of magnesium from EF-Tu nucleotide complexes with EDTA results in a loss in affinity for the nucleotide. It has also been suggested that guanine-nucleotide exchange factors might reduce the affinity for magnesium ions<sup>21</sup>. Based on the changes in the vicinity of the nucleotide binding site of EF-Tu (Fig. 4c, d), the mechanism for the release of nucleotides catalysed by EF-Ts appears to be predominantly due to disruption of the magnesium binding site in domain 1 of EF-Tu, which is induced by the intrusive interactions of sAsp 80 and sPhe 81. This leads to the displacement of helix B and the strand containing uAsp 80, uCys 81 and uPro 82, all of which are involved in the stabilization of water molecules coordinating the magnesium ion in the EF-Tu·GDP complex.

A secondary factor contributing to nucleotide release is a subtle but important alteration in the vicinity of the  $\beta$ -phosphate binding site. While the interaction of uHis 19 with the peptide oxygen of sGly 126 would tend to shift the phosphate binding loop, the interaction of uAsp 21 and sLys 51 would maintain it in place. Thus the first four amino acid residues of the loop (and the  $\beta$ -strand upstream) are displaced, but the last four are not, flipping the peptide bond between uVal 20 and uAsp 21. This peptide flip would break the hydrogen bond between the  $\beta$ -phosphate oxygen and the peptide nitrogen of uAsp 21 observed in the EF-Tu·DGP structure<sup>8</sup>, and would place the peptide oxygen of uVal 20 in a position to make phosphate binding no longer energetically favourable. Note that in the GDNPP bound structure<sup>10</sup>, the  $\beta$ -phosphate adopts a conformation such that interference with the above peptide oxygen would not occur, so it would not prevent

GTP binding, to form a ternary GTP·EF-Tu·EF-Ts complex. The last significant contribution to the nucleotide release is the movement of helix D, which shifts uLys 136 and uAsp 138, involved in the stabilization of the ribose and the guanine base, away from the nucleotide binding site, thereby relaxing the interactions with the base.

### Discussion

The structure of the EF-Tu·EF-Ts complex presented above, in which the EF-Ts component is a tight dimer, raises the question of stoichiometry in the classical model for the elongation cycle in bacterial protein biosynthesis. In this context, it should also be noted that in eukaryotic cytosols the polypeptide elongation factors constitute a multimeric complex, EF-1, and in *Artemia* this complex is pentameric with a stoichiometry of 2:1:1:1 for the  $\alpha$ : $\beta$ : $\gamma$ : $\delta$  subunits<sup>22</sup>, where EF-1 $\alpha$  is the homologue of EF-Tu, and the two factors EF-1 $\beta$  and EF-1 $\delta$  are nucleotide exchange factors. Related to this question of stoichiometry is the quarternary structure of the bacteriophage RNA polymerase of Q $\beta$ -replicase: this contains EF-Tu, EF-Ts, the ribosomal protein S1, and a bacteriophage coded protein in equimolar amounts; estimations of the molecular weight of the replicase indicate that only one copy of each subunit was present in the complex<sup>23</sup>.

The interaction of EF-Ts with domain 1 of EF-Tu indicates that the driving force for the exchange of guanine nucleotides is the insertion of the amino acid side chains of a conserved peptide sequence, TDFV, into the EF-Tu structure. This disrupts the binding site for Mg<sup>2+</sup>, leading to a loss of affinity for the nucleotide. The critical interactions appear to be between the hydrophobic side chain of phenylalanine and the carboxyl group of aspartic acid with EF-Tu. It is pertinent to ask whether this mechanism also applies to other guanine-nucleotide exchange factors and, if so, whether corresponding sequence motifs can be identified. In the eukaryotic exchange factors EF-1 $\beta$  and EF-1 $\delta$ , a consensus sequence, ETDM, can be found which may correspond to the EF-Ts ETDFV consensus sequence. Furthermore, for rasGEF, the sequence of the SCR p-0 region<sup>24</sup> and an extension of 10 residues downstream can be aligned with the TDM part of the eukaryotic motif and downstream residues. □

Received 25 October; accepted 19 December 1995.

1. Ertel, R., Redfield, B., Brot, V. & Weissbach, H. *Archs Biochem. Biophys.* **128**, 331–338 (1968).
2. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117–127 (1991).
3. Lucus-Lenard, J. & Lipmann, F. *Proc. natn. Acad. Sci. U.S.A.* **55**, 1562–1566 (1966).
4. Weissbach, H., Miller, D. L. & Hachmann, J. *Archs Biochem. Biophys.* **137**, 262–269 (1970).
5. Miller, D. L. & Weissbach, H. *Archs Biochem. Biophys.* **141**, 26–37 (1970).
6. Leberman, R. et al. *Analyt. Biochem.* **104**, 29–36 (1980).
7. Blumental, T., Landers, T. A. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1313–1317 (1972).
8. Kjeldgaard, M. & Nyborg, J. *J. molec. Biol.* **223**, 721–742 (1992).
9. Kabsch, W., Gast, W. H., Schulz, G. E. & Leberman, R. *J. molec. Biol.* **117**, 999–1012 (1977).
10. Pal, E. F. et al. *Nature* **341**, 209–214 (1989).
11. Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. *Structure* **1**, 35–50 (1993).
12. Berchtold, H. et al. *Nature* **365**, 126–132 (1993).
13. Jones, M. D. et al. *Eur. J. Biochem.* **108**, 507–526 (1980).
14. An, G., Bendiak, D. S., Mamelak, L. A. & Friesen, J. D. *Nucleic Acids Res.* **9**, 4163–4172 (1981).
15. Leberman, R., Schulz, G. E. & Suck, D. *FEBS Lett.* **124**, 279–281 (1981).
16. Schneider, U. dissertation, Univ. Freiburg, Germany (1983).
17. Schick, B. & Jurnak, F. *Acta crystallogr.* **D50**, 563–568 (1994).
18. Kawashima, T., Berthet-Colominas, C., Cusack, S. & Leberman, R. *Acta crystallogr.* (in the press).

19. Collaborative Computing Project No. 4 (SERC Daresbury Laboratory, Warrington, 1979).
20. Arai, K.-I., Kawakita, M. & Kaziro, Y. *J. biol. Chem.* **247**, 7029–7037 (1972).
21. John, J. et al. *J. biol. Chem.* **268**, 923–929 (1993).
22. Blumental, T. & Carmichael, G. C. A. *Rev. Biochem.* **48**, 525–548 (1979).
23. Janssen, G. M. C., van Damme, H. T. F., Kriek, J., Amons, R. & Möller, W. *J. biol. Chem.* **269**, 32410–32417 (1994).
24. Boguski, M. S. & McCormick, G. *Nature* **366**, 643–654 (1993).
25. Leslie, A. G. W. in *Proceedings of the CCP4 Study Weekend* (eds Helliwell, J. R., Machin, P. A. & Papiz, M. Z.) 39–50 (SERC Daresbury Laboratory, Warrington, 1987).
26. Jones, T. A., Cowan, S., Zou, J.-Y. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1990).
27. Jones, T. A. in *Molecular Replacement* (eds Dodson, E. J., Gover, S. & Wolf, W.) 92–105 (SERC Daresbury Laboratory, Warrington, 1992).
28. Brunger, A. T. *Nature* **355**, 472–475 (1992).
29. Kraulis, P. J. *appl. Crystallogr.* **24**, 946–950 (1991).
30. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).

SUPPLEMENTARY INFORMATION: Available on Nature's World-Wide Web site <http://www.nature.com>; requests may also be addressed to Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank O. Wiborg for the EF-Ts clone; J. Nyborg for the coordinates of the EF-Tu·GDP structure; and P. B. Sigler, M. Sprinzl and A. Wittinghofer for discussions.

# A nearby massive galaxy cluster behind the Milky Way

R. C. Kraan-Korteweg<sup>\*†</sup>, P. A. Woudt<sup>‡</sup>, V. Cayatte<sup>\*</sup>,  
A. P. Fairall<sup>‡</sup>, C. Balkowski<sup>\*</sup> & P. A. Henning<sup>§</sup>

<sup>\*</sup> Observatoire de Paris, DAEC, Unité associée au CNRS, D0173,  
et à l'Université Paris 7, 92195 Meudon Cedex, France

<sup>‡</sup> Department of Astronomy, University of Cape Town, Rondebosch,  
7700 South Africa

<sup>§</sup> Department of Physics and Astronomy, University of New Mexico,  
800 Yale Boulevard Northern, Albuquerque, New Mexico 87131, USA

THE recession velocities of relatively nearby galaxies show systematic deviations from a uniform expansion field. These deviations indicate the presence of the 'Great Attractor'—a large concentration of mass ( $\sim 5 \times 10^{16}$  solar masses) that lies in the direction of the southern Milky Way<sup>1,2</sup>. Attempts to quantify both the nature and extent of the Great Attractor<sup>3–11</sup> have been hampered by the fact that it is largely hidden by the disk of the Milky Way. Although there is an excess of galaxies in this region<sup>12</sup>, no dominant galaxy cluster or other concentration of mass has hitherto been identified. Here we present results from a survey of galaxies obscured (or partially obscured) by the southern Milky Way. Our results show that a previously identified galaxy cluster, Abell 3627, which lies only  $9^\circ$  from the predicted centre of the Great Attractor<sup>13</sup>, is very massive ( $\sim 5 \times 10^{15}$  solar masses). The cluster's redshift is also near that predicted for the core of the Great Attractor<sup>2,3,12–14</sup>, suggesting that it lies at or near the bottom of the Attractor's gravitational potential well.

Finding the nature and extent of the Great Attractor has been the aim of many optical<sup>3–6</sup> and IRAS galaxy surveys<sup>7–10</sup>, as well as X-ray searches<sup>11</sup>. Although the arguments<sup>2,7</sup> for its being a nearby inhomogeneity (recession velocity  $v \approx 4,500 \text{ km s}^{-1}$ ) have been challenged by those who look to much larger scales<sup>15,16</sup>, recent consensus<sup>12,14</sup> is that it is probably an extended region of moderately increased density. Although there is a considerable excess of optical or IRAS galaxies<sup>12</sup> in this region, no dominant cluster or central peak has been identified.

The location of the region's centre is poorly defined because of measurement errors in the observations, scatter and zero-point uncertainty in the adopted distance relations, and the large smoothing of the sparsely sampled field. Calculations<sup>13</sup> with the reconstruction algorithm POTENT, which maps the underlying density field from velocity flow fields<sup>17</sup> based on a large data set of elliptical and spiral galaxies, now put the centre of the potential field at galactic longitude and latitude  $l \approx 320^\circ$ ,  $b \approx 0^\circ$ . This lies in the middle of the Milky Way.

The dust and gas in our Galaxy obscure the light of the galaxies behind its disk, creating a 'Zone of Avoidance' (ZOA) in their distribution. Even if the core of the Great Attractor were in the form of luminous matter, it would be obscured.

In an attempt to gauge the luminous mass distribution in the ZOA, we have for some years been searching<sup>18,19</sup> for partially obscured galaxies behind the southern Milky Way ( $|b| \lesssim 10^\circ$ ) by visual examination of existing sky surveys. Our survey covers much of the Great Attractor region. The portion of sky concerned is shown in Fig. 1 with a  $40^\circ \times 40^\circ$  field, roughly the diameter of the density enhancement. Previously identified larger galaxies (diameter  $D \geq 1.0'$ )<sup>20</sup> are shown as blue dots. Red dots mark galaxies identified by us in our search for galaxies down to a diameter limit of  $D = 0.2'$ . The low diameter limit was chosen because of obscuration effects. For an extinction of the blue light of  $A_B = 2.5 \text{ mag}$  the apparent angular diameter will be reduced by about a factor of 3 and the galaxy images will fade altogether if

$A_B \gtrsim 5\text{--}6 \text{ mag}$  (refs 21, 22). This is seen directly in the galaxy distribution shown in Fig. 1. The superposed contours indicate the distribution of the Galactic hydrogen gas, which is a good tracer of the foreground extinction if the gas-to-dust ratio is constant. The innermost part of the Milky Way remains fully opaque but the ZOA has been reduced to a strip  $5\text{--}6^\circ$  wide. Distinct extragalactic large-scale features can be recognized outside this band.

Figure 1 is centred on the predicted direction of the Great Attractor. The previously known larger galaxies (blue dots) show little concentration towards the centre, save for a general enhancement running from bottom left to top right. Similarly, IRAS galaxy candidates taken from the extension to low latitudes of the Point Source Catalog redshift-survey<sup>23</sup>, (not plotted here), show a smooth distribution, with no condensation towards the Great Attractor. Our survey (red dots) however, reveals a marked overdensity—at least a factor of 10 denser than regions at similar extinctions—at a short distance southeast of the centre of the Great Attractor.

This overdensity is centred on a previously identified cluster, A3627. The cluster is listed in the extended whole-sky Abell catalogue<sup>24</sup> which contains 4,073 'rich' galaxy clusters: that is, clusters with more than 50 galaxies ( $N_A \geq 50$ ) in the magnitude range between the third brightest galaxy  $m_3$  and  $m_3 + 2$  found within the Abell radius  $R_A = 1.7/z$ , where  $z$  is the redshift ( $v_{\text{obs}}/c$ ). Although it is the only Abell cluster identified behind the Milky Way, classified as a nearby, rich cluster ( $N_A$  is about 59), and close to the expected position of the Great Attractor, this cluster has received little attention<sup>25</sup>. Part of the reason is the effect of the foreground obscuration. This cluster at  $(l, b) = (325^\circ, -7^\circ)$  is

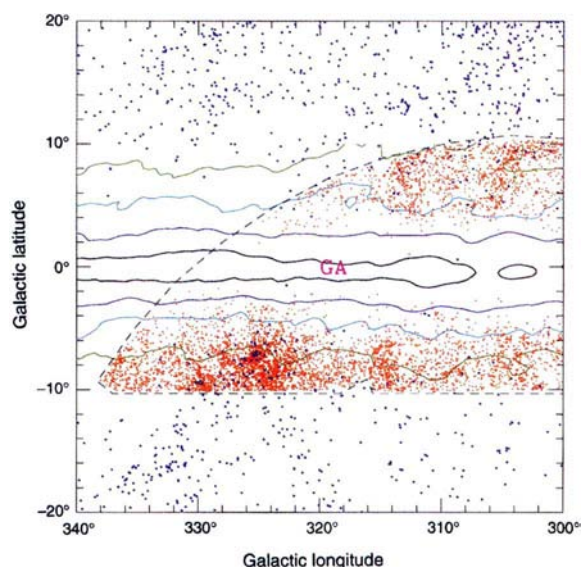


FIG. 1 Area of  $40^\circ \times 40^\circ$  centred on the predicted position of the Great Attractor (GA). The blue dots mark the Lauberts galaxies (diameter  $D \geq 1.0'$ ), the red dots the partially obscured galaxies ( $D \geq 0.2'$ ) identified through our visual inspection of sky survey plates. The dashed lines demarcate the area surveyed so far. As an approximation of the foreground extinction, contours marking the Galactic distribution of neutral hydrogen gas (H I) are superposed. The green, light-blue, blue and black contours mark H I column densities<sup>38</sup> of 2.2, 3.7, 6.4 and  $14.6 \times 10^{21} \text{ atom cm}^{-2}$ , or extinctions of  $A_B = 1.5, 2.7, 4.8, 11.3 \text{ mag}$ . Although it can be argued that the adopted transformation<sup>39</sup> overestimates the extinction in this part of the Galactic plane by a factor of two<sup>40</sup>, this will not change the general extinction pattern. Note that the Milky Way is nearly fully opaque below the blue contour, whereas distinct extragalactic large-scale structures are recognized above the light-blue contour, the most prominent being an overdensity centred on  $l = 325^\circ$ ,  $b = -7^\circ$ , some  $9^\circ$  from the position of the Great Attractor.

<sup>†</sup> Formerly at Kapteyn Astronomical Institute, PO Box 800, 9700 AV Groningen, The Netherlands.



hardly obvious in the distribution of Lauberts galaxies (blue dots in Fig. 1). But the distribution of galaxy diameters of  $D = 1.0'$  in this overdensity peaks just below the diameter limit of Lauberts. In addition, the magnitudes ( $B_J$ ) and diameters ( $D$ ) are significantly brighter and larger than the mean of our survey ( $B_J = 16.7$  mag,  $D = 0.7'$  compared with  $B_J = 18.0$  mag,  $D = 0.4'$ ), or the mean at similar latitudes. Lifting the obscuring veil of the Milky Way by correcting the observed properties of the galaxies for the foreground extinction would reveal 139 galaxies with  $D \geq 1.0'$  in the central part of the cluster ( $R_A \leq 1.75^\circ$ ) rather than the 31 galaxies previously identified<sup>20</sup>, and reveal this cluster as the most prominent overdensity in the southern sky.

To assess the implications of this and other newly discovered extragalactic features for the dynamics of the nearby Universe, we are measuring the redshifts of the galaxies uncovered in our deep survey with three complementary observational approaches<sup>26</sup>: multifibre spectroscopy with MEFOS at the 3.6-m telescope of European Southern Observatories (ESO) in the densest regions; individual spectroscopy of the brightest galaxies with the 1.9-m telescope of the South African Astronomical Observatory (SAAO); and 21-cm observations of extended, low-surface-brightness spirals with the 64-m radio telescope in Parkes, Australia. With the resulting redshift coverage (typically  $\geq 10\%$  of the uncovered galaxies) we can trace large-scale structures out to recession velocities of  $v \geq 25,000 \text{ km s}^{-1}$ .

Although the cluster includes the well researched radiogalaxy PKS1610 – 60 (ref. 27), few redshifts of other cluster galaxies are available from the literature<sup>28</sup>. Moreover, when plotted as a histogram (dark shading in Fig. 2), these redshifts are too scattered to allow interpretation. But when one adds our 95 new velocities (lighter shading in Fig. 2) within the Abell radius  $R_A = 1.75^\circ$  and a velocity range,  $1,000 \leq v_{\text{obs}} \leq 9,000 \text{ km s}^{-1}$ , this reveals a distribution close to pure gaussian (an 'a-test'<sup>29</sup> gives a value of 0.807, compared to a value of 0.798 for an infinite gaussian). The plot shows almost negligible foreground and background contamination—it happens that there are voids in front of and behind the cluster. The mean observed velocity and dispersion of the cluster are  $\langle v_{\text{obs}} \rangle = 4,882 \text{ km s}^{-1}$  (or  $4,655 \text{ km s}^{-1}$  with respect to the centroid of the Local Group of galaxies) and  $\sigma = 903 \text{ km s}^{-1}$  ( $\sigma = 897 \text{ km s}^{-1}$  when corrected for measurement errors). This puts the cluster well within the predicted velocity range of the Great Attractor.

The large dispersion of A3627 suggests that it is fairly massive. The mass,  $M$ , can be determined from the virial theorem using the formalism<sup>30</sup>  $M = 9R_{\text{eff}}\sigma^2/G$ , if the logarithmic distribution of the surface densities follows an  $R^{1/4}$  profile where  $R$  is the distance from the centre of the cluster, and the effective radius  $R_{\text{eff}}$  a scale

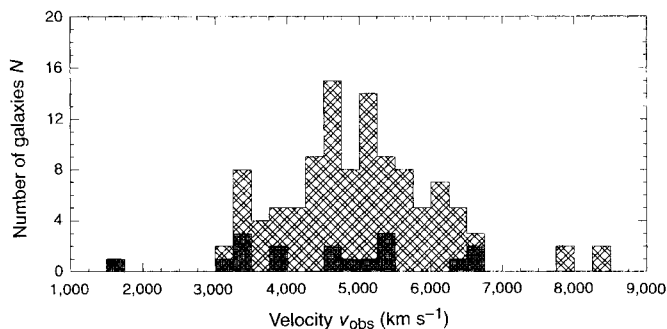


FIG. 2 Velocity distribution of the 112 galaxies within the Abell radius  $R_A = 1.75^\circ$  of A3627, for which redshifts in the displayed velocity range have been determined. Dark shading indicates previously published redshifts, lighter shading shows our new measurements. The mean velocity and dispersion of the cluster are  $\langle v_{\text{obs}} \rangle = 4,882 \text{ km s}^{-1}$ ,  $\sigma = 903 \text{ km s}^{-1}$ , not including the one galaxy at  $\sim 1,500 \text{ km s}^{-1}$  and the four galaxies at  $\sim 8,000 \text{ km s}^{-1}$ .

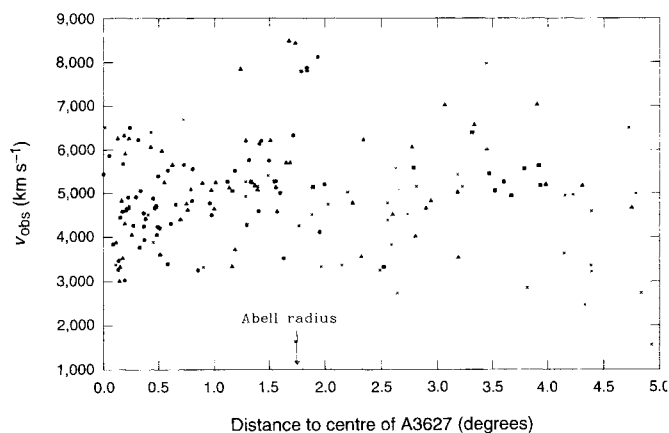


FIG. 3 Distribution of the 131 new velocities between  $1,000 \leq v_{\text{obs}} \leq 9,000 \text{ km s}^{-1}$  in a circle of radius  $5^\circ$  around A3627, as a function of distance from the centre. The symbols identify the three different instruments used: MEFOS multifibre spectroscopy at the 3.6-m ESO telescope (dots,  $N = 61$ ), individual spectroscopy at the 1.9-m SAAO telescope (triangles,  $N = 58$ ) and H I detections with the Parkes 64-m radiotelescope (boxes,  $N = 12$ ). The additional 43 crosses are previously measured redshifts from the literature. With the exception of a small group at  $8,000 \text{ km s}^{-1}$ , the cluster is well isolated from the foreground and background. But note that the mean velocity and velocity dispersion do not change within the  $5^\circ$  circle shown here, suggesting that A3627 is embedded in a much larger structure.

parameter). This is indeed the case. Because of possible background contamination, however, we restrict the calculation of the mass to the inner 3 Mpc, the radius within which most clusters are virialized, rather than  $R_{\text{eff}} = 155' = 4.2h_{50}^{-1} \text{ Mpc}$ . This results in a mass for A3627 of  $M_{\text{A3627}} = 5.1 \times 10^{15} h_{50}^{-1} M_\odot$ , where  $M_\odot$  is solar mass. The mass is similar to that of the Coma cluster<sup>31,32</sup> and typical for rich clusters.

The richness of A3627 is apparent from the density and the predominance of early-type galaxies in its core. The core radius (defined as the radius where density has fallen to half the central value) is  $R_c = 10.4'$  with a central density of  $N_0 = 800$  galaxies per square degree. Rich massive clusters are generally strong X-ray emitters and were mostly identified early on by X-ray satellites (such as Einstein, HEAO and Uhuru). A3627 has so far never been identified in X-ray<sup>11,33</sup>, unlike, for instance, the nearby rich clusters Perseus and Coma. However, recent inspection of the Rosat data find A3627 to be the sixth brightest galaxy cluster in the Rosat band (0.1–2.4 keV) and Rosat observations with the Position Sensitive Proportional Counter confirm it to be as massive as Coma<sup>41</sup>. Moreover, the independent mass estimate from the X-ray data proves the virial mass to be the true mass.

The derived properties of A3627—its velocity dispersion, central density and core radius ( $R_c = 0.3 h_{50}^{-1} \text{ Mpc}$  for a distance of  $d \approx v_0/H_0 = 93 h_{50}^{-1} \text{ Mpc}$  where  $H_0$  is the Hubble constant), its large fraction of early-galaxies (50% in the core of the cluster), mass and Rosat X-ray luminosity ( $2.2 \times 10^{44} \text{ erg s}^{-1}$ ) are all consistent with its being the nearest known rich massive cluster in the Universe. In fact, simulations show that if Coma were at the position of A3627 in velocity space and subjected to the same foreground obscuration the two clusters would be indistinguishable. A3627 even has, like Coma, two dominant cD galaxies at its core.

Moreover, the location of A3627 within the overdensity associated with the Great Attractor makes this massive cluster the prime candidate for being at the centre of the region. If the Great Attractor is an extended region of moderately increased density, then biasing would predict the existence of a rich cluster near such a large-scale potential well<sup>13</sup>. This is consistent with our finding. A3627 is indeed embedded in a larger structure. As illustrated in

Fig. 3, which resembles that for the Coma cluster<sup>34</sup>, A3627 can be traced well beyond its Abell radius of  $R_A = 1.75^\circ$ . The mean redshift remains roughly constant to at least  $5^\circ$  from the centre of the cluster. The results from our programme across the whole search area suggest furthermore that A3627 is part of a super-cluster, starting at the nearby Pavo, Indus clusters below the Galactic plane, moving across A3627 towards the shallow overdensity in Vela at  $6,000 \text{ km s}^{-1}$  discovered earlier<sup>35,36</sup>. Three-dimensional visualizations in redshift space of the galaxy distribution combining our most recent redshift data with the distribution of all known galaxies in the Southern Hemisphere<sup>28,37</sup> provide further support for the idea that A3627 is the dominant component of a 'great wall' structure, similar to the Coma cluster in the (northern) Great Wall.  $\square$

Received 22 June; accepted 19 December 1995.

- Dressler, A. et al. *Astrophys. J.* **313**, L37–L42 (1987).
- Lynden-Bell, D. et al. *Astrophys. J.* **326**, 19–49 (1988).
- Dressler, A. *Astrophys. J.* **329**, 519–526 (1988).
- Hudson, M. *Mon. Not. R. astr. Soc.* **265**, 43–71 (1993).
- Hudson, M. *Mon. Not. R. astr. Soc.* **265**, 72–80 (1993).
- Rowan-Robinson, M. et al. *Mon. Not. R. astr. Soc.* **247**, 1–18 (1990).
- Strauss, M. A., Davis, M., Yahil, A. & Huchra, J. P. *Astrophys. J.* **361**, 49–62 (1990).
- Strauss, M. A., Yahil, A., Davis, M., Huchra, J. P. & Tonry, J. *Astrophys. J.* **397**, 395–419 (1992).
- Hoffman, Y. in *Cosmic Velocity Fields, Proc. 9th IAP Astrophys. Mtg* (eds Bouchet, F. R. & Lachièze-Rey, M.) 357–366 (Editions Frontières, Gif-sur-Yvette, 1994).
- Jahoda, K. & Mushotzky, R. F. *Astrophys. J.* **346**, 638–647 (1989).
- Lynden-Bell, D. in *Observational Test of Cosmological Inflation* (eds Shanks, T. et al.) 337–353 (Kluwer, Dordrecht, 1991).
- Kolatt, T., Dekel, A. & Lahav, O. *Mon. Not. R. astr. Soc.* **275**, 797–811 (1995).
- Hudson, M. *Mon. Not. R. astr. Soc.* **266**, 475–488 (1994).
- Scaramella, R., Vettolani, G. & Zamorani, G. *Astrophys. J.* **376**, L1–L4 (1991).
- Lauer, T. & Postman, M. *Astrophys. J.* **425**, 418–438 (1994).
- Dekel, A. A. *Rev. Astr. Astrophys.* **32**, 371–418 (1994).
- Kraan-Korteweg, R. C. in *Rev. mod. Astr.* **2**, 119–128 (1989).
- Kraan-Korteweg, R. C. & Woudt, P. A. in *Unveiling Large-Scale Structures behind the Milky Way*,

- Proc. 4th DAEC Mtg* (eds Balkowski, C. & Kraan-Korteweg, R. C.) 89–98 (Astr. Soc. Pacific Conf. Ser. 67, San Francisco, 1994).
- Lauberts, A. *The ESO/Uppsala Survey of the ESO (B) Atlas* (ESO, Garching, 1982).
- Kraan-Korteweg, R. C. in *The Distribution of Matter in the Universe, Proc. 2nd DAEC Workshop* (eds Mamon, G. & Gerbal, D.) 202–211 (Observatoire de Paris, Paris, 1993).
- Cameron, L. M. *Astr. Astrophys.* **233**, 16–20 (1990).
- Saunders, W. et al. in *Unveiling Large-Scale Structures behind the Milky Way, Proc. 4th DAEC Mtg* (eds Balkowski, C. & Kraan-Korteweg, R. C.) 257–267 (Astr. Soc. Pacific Conf. Ser. 67, San Francisco, 1994).
- Abell, G. O., Corwin, H. G. & Olowin, R. P. *Astrophys. J. Suppl. Ser.* **70**, 1–138 (1989).
- Mould, J. R. et al. *Astrophys. J.* **383**, 467–486 (1991).
- Kraan-Korteweg, R. C., Cayatte, V., Fairall, A. P., Balkowski, C. & Henning, P. A. in *Unveiling Large-Scale Structures behind the Milky Way, Proc. 4th DAEC Mtg* (eds Balkowski, C. & Kraan-Korteweg, R. C.) 99–111 (Astr. Soc. Pacific Conf. Ser. 67, San Francisco, 1994).
- Danziger, I. J. & Goss, W. M. *Mon. Not. R. astr. Soc.* **202**, 703–715 (1983).
- Fairall, A. P. in *Unveiling Large-Scale Structures behind the Milky Way, Proc. 4th DAEC Mtg* (eds Balkowski, C. & Kraan-Korteweg, R. C.) 77–78 (Astr. Soc. Pacific Conf. Ser. 67, San Francisco, 1994).
- Yahil, A. & Vidal, N. V. *Astrophys. J.* **214**, 347–350 (1977).
- Rood, H. J., Page, T. L., Kintner, E. C. & King, I. R. *Astrophys. J.* **175**, 627–647 (1972).
- Hughes, J. P. *Astrophys. J.* **337**, 21–33 (1990).
- White, S. D. M., Briel, U. G. & Henry, J. P. *Mon. Not. R. astr. Soc.* **261**, L8–L12 (1993).
- Lahav, O., Edge, A. C., Fabian, A. C. & Putney, A. *Mon. Not. R. astr. Soc.* **238**, 881–895 (1989).
- Kent, S. M. & Gunn, J. E. *Astr. J.* **87**, 945–971 (1982).
- Kraan-Korteweg, R. C., Fairall, A. P. & Balkowski, C. *Astr. Astrophys.* **297**, 617–633 (1995).
- Kraan-Korteweg, R. C. & Woudt, P. A. in *Proc. 9th IAP Astrophysics Mtg on Cosmic Velocity Fields* (eds Bouchet, F. & Lachièze-Rey, M.) 557–558 (Editions Frontières, Gif-sur-Yvette, 1994).
- Fairall, A. P. & Paverd, W. R. *Wide Field Spectroscopy and the Distant Universe* (eds Maddox, S. J. & Aragon-Salamanca, A.) 121–126 (World Scientific, London, 1995).
- Kerr, F. J., Bowers, P. F., Jackson, P. D. & Kerr, M. *Astr. Astrophys. Suppl. Ser.* **66**, 373–504 (1986).
- Burstein, D. & Heiles, C. *Astr. J.* **87**, 1165–1189 (1982).
- Burstein, D. et al. *Astrophys. J. Suppl. Ser.* **64**, 601–647 (1987).
- Bohringer, H., Neumann, D. M., Schindler, S. & Kraan-Korteweg, R. C. *Astrophys. J.* (in the press).

**ACKNOWLEDGEMENTS.** We thank H. Böhringer for sharing his early ROSAT results on the cluster A3627, H. Bloemen for providing integrated H I data derived from observations, the MEPOS team for assistance, the staff at the SAAO and at Parkes Telescope for support and hospitality, B. Bassett for assistance at SAAO and A. Yahil for constructive feedback. R.C.K.K. was supported by an EC grant; research in Groningen was supported by a fellowship of the Royal Netherlands Academy of Arts and Sciences. Support was also provided by CNRS through the Cosmology GDR programme. A.P.F. and P.A.W. are supported by the South African FRD.

## Equivalence of the sound velocity in water and ice at mesoscopic wavelengths

G. Ruocco\*, F. Sette†, U. Bergmann†, M. Krisch†, C. Masciovecchio†, V. Mazzacurati\*, G. Signorelli\* & R. Verbeni†

\* Università di L'Aquila and Istituto Nazionale di Fisica della Materia, I-67100, L'Aquila, Italy

† European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble Cedex, France

It is generally assumed, for any material, that sound travels faster in the solid phase than in the liquid. This is true at least for waves of macroscopic wavelength, the dynamics of which depend on the elastic properties of the medium and, hence, on the presence or absence of long-range order. But at wavelengths approaching the typical interparticle distance, the dynamics should become insensitive to the long-range organization of the medium. Here we report inelastic X-ray scattering measurements in water and ice, which show that sound waves with wavelengths between 0.5 and 3 nm propagate at the same velocity in both phases. The observed sound speed, which is greater than twice the hydrodynamic speed of sound in water, but less than that in ice, agrees with values obtained in previous measurements of 'fast sound' in liquid D<sub>2</sub>O (ref. 1) and H<sub>2</sub>O (ref. 2). These results show that, despite the fundamental structural and dynamical differences between water and ice, the dynamical response of the two phases is strikingly similar at very short wavelengths.

The liquid and solid states are thermodynamical concepts related to the time evolution of the relative positions of the

constituents of a condensed material. Whereas in the solid there are no changes over a long timescale, the liquid has characteristic diffusive motions. These two concepts are at the basis of many macroscopic properties of materials, often observable in everyday life. In contrast, in microscopic processes such as local bonding, first-neighbour atomic structure and interatomic interactions, the differences between the liquid and the solid phase often become irrelevant to the physical problems under investigation. In the intermediate, mesoscopic, time-space domain, characteristic of important processes such as bond formations, chemical reactions and aggregation of complex systems, density fluctuations in the liquid and solid phases may deviate from macroscopic behaviour, although they still involve many atoms.

In general, the dynamics of density fluctuations can be assessed experimentally by determining the dynamic structure factor,  $S(Q, \omega)$ , which is the time-and-space Fourier transform of the density-density correlation function<sup>3</sup>.  $S(Q, \omega)$  is measured by inelastic scattering experiments, where  $\hbar Q$  and  $\hbar \omega$  are the momentum and energy transferred to the system by the probe particle. Features in  $S(Q, \omega)$  are directly related to the existence of elementary excitations with wavelength  $2\pi/Q$ , and frequency  $\omega$ . In the macroscopic time-space domain, these excitations, known as sound-waves, have a linear dispersion between  $\omega(Q)$  and  $Q$ , and propagate with a characteristic velocity given by the ratio  $\omega(Q)/Q$ . In the mesoscopic time-space domain, corresponding to a typical energy-momentum region of 0.1–20 meV and  $10^{-2}$ – $10 \text{ nm}^{-1}$ , the existence of propagating sound-like excitations in the liquid is questionable.

Measurements of  $S(Q, \omega)$  in this range are traditionally done using inelastic neutron scattering. But because of the kinematics of the scattering process, this technique cannot be easily applied to materials in which the speed of sound is greater than  $1,500 \text{ m s}^{-1}$ . In crystalline systems, this limitation is overcome by the periodicity of the lattice, which allows researchers to use neutrons to study dynamics in higher Brillouin zones. In liquids and disordered

materials, where there is no translational invariance, collective excitations can be studied only at momentum transfers smaller than the inverse of the interparticle distances. This explains the very small number of neutron studies of liquids, and the absence of a direct comparison between the high-frequency collective dynamics of a material below and above the melting transition.

The kinematic limitations of neutrons do not apply to inelastic X-ray scattering with 10–20 keV incident energy, but the technique has not often been used, as a consequence of poor energy resolution and photon flux. Developments in perfect crystal optics, and the advent of third-generation synchrotron radiation sources, have allowed us to overcome these difficulties. We have measured  $S(Q, \omega)$  for water by inelastic X-ray scattering with millielectronvolt energy resolution at the new Inelastic X-ray Scattering Beamline (BL21-ID16) at the European Synchrotron Radiation Facility. The X-ray source is an undulator used in fifth harmonic mode to produce X-rays of about 18 keV. The X-ray beam is monochromatized using the Si(999) reflection from a perfect silicon crystal (intrinsic relative energy resolution of  $1.1 \times 10^{-7}$ ) positioned in extreme backscattering geometry to maximize the flux<sup>4–7</sup>. This monochromatic beam has an energy of 17,794 eV, an energy resolution of 2 meV and an intensity of  $10^9$  photons  $s^{-1}$ , and is focused on the sample to a spot size of  $150 \times 350 \mu m^2$  by a toroidal mirror. The scattered photons are collected by a spherical silicon crystal analyser with 7 m radius, operating also at the Si(999) in backscattering<sup>8</sup>.

The total energy resolution function was experimentally determined, and it is well represented by a lorentzian profile with 3.2 meV full width at half maximum (FWHM). The scattering angle for momentum transfers between 2 and  $14 \text{ nm}^{-1}$  was selected by rotating the 7-m analyser arm in the horizontal scattering plane. The  $Q$  resolution was set to  $0.4 \text{ nm}^{-1}$  by an aperture in front of the analyser. Energy scans were done by

varying the relative temperature between the monochromator and analyser crystals with millikelvin precision. Further experimental details on the beamline are given elsewhere<sup>2,9</sup>. The water sample was high-purity  $H_2O$  kept at  $4^\circ \text{C}$  in the liquid measurements, and at  $-20^\circ \text{C}$  in the solid. The ice samples were in the 1h crystal structure, characterized by an ordered hexagonal lattice of oxygen atoms and an almost random distribution of the hydrogen atoms on one of the two equivalent positions along the nearest-neighbour O–O directions. The dispersion was measured in a polycrystal, and in a single crystal along the (1011) and (0001) crystallographic directions. No relevant differences were observed at the same absolute  $Q$  value, in agreement with lattice dynamics calculations<sup>10</sup>. The sample thicknesses were  $\sim 15 \text{ mm}$  in the direction of the incident beam.

Inelastic X-ray scattering spectra at selected  $Q$ -values and the corresponding fits are shown in Fig. 1. In Fig. 1a we show the spectra of polycrystalline ice. At low  $Q$  the spectra are dominated by the longitudinal acoustic (LA) phonon mode whose energy increases with  $Q$ . The non-dispersive peak at  $\sim 7 \text{ meV}$ , appearing at  $Q$  larger than  $6 \text{ nm}^{-1}$ , is the transverse optic phonon mode (TO) which in this  $Q$  region acquires some longitudinal character. In Fig. 1b we show the spectra of liquid water at the indicated  $Q$  transfers. In the liquid, the spectra are dominated by the quasi-elastic central peak. The curve aligned with this peak is the measured resolution function, which we show to emphasize the presence of scattered intensity at the sides of the central peaks. This inelastic intensity is due to collective excitations in water. To determine the energy positions of the excitations in Fig. 1, we fitted the spectra as explained in the figure legend.

In Fig. 2 we show the dispersion of the longitudinal collective excitations for liquid and solid water as obtained from the fitting procedure. The reported energies correspond to the maxima of the longitudinal current spectra<sup>11</sup>. In the crystal, this is the

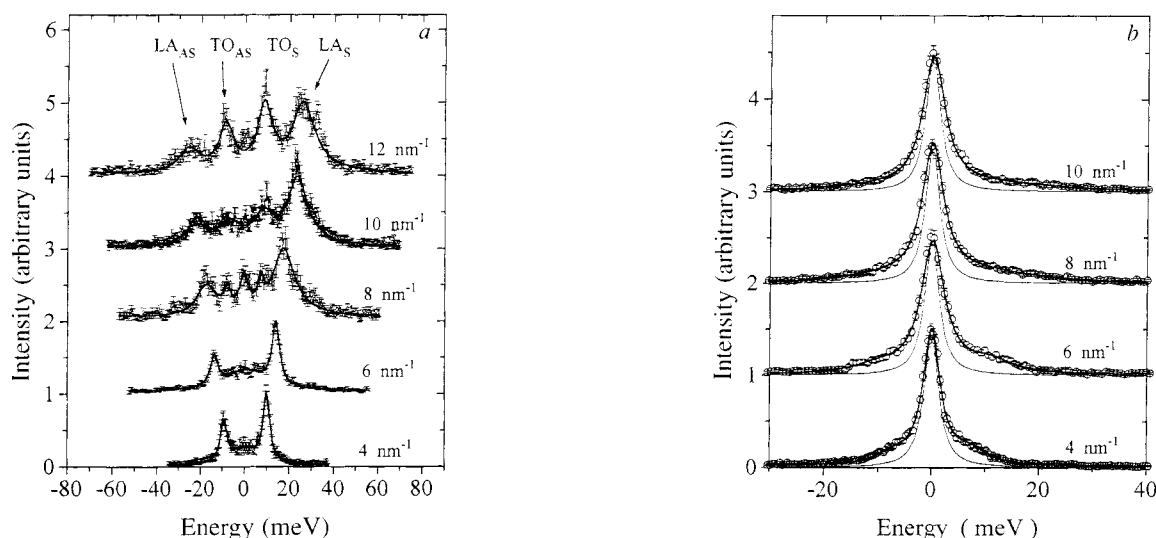


FIG. 1 *a*, Inelastic X-ray scattering spectra of polycrystalline  $H_2O$  ice 1h at  $-20^\circ \text{C}$ , taken at the indicated values of  $Q$ . The longitudinal acoustic and the transverse optical phonons are labelled LA and TO, respectively. The experimental data are normalized to the maximum of the LA Stokes peak. At this point, the count rates were  $\sim 1 \text{ count s}^{-1}$  at all the  $Q$  values investigated. The integration time for each data point was 180 s at  $Q = 4$  and  $6 \text{ nm}^{-1}$  and 90 s at  $Q = 8, 10$  and  $12 \text{ nm}^{-1}$ . The data points ( $\circ$ ), shown with error bars ( $\pm \sigma$ ), are superimposed on the fit (solid line). The fit was made by the convolution of the experimental resolution function with two pairs of lorentzians, representing the LA and TO Stokes (S) and anti-Stokes (AS) peaks. A fifth lorentzian was used to account for the small elastic intensity at  $Q = 4, 6$  and  $8 \text{ nm}^{-1}$ . At  $Q = 10$  and  $12 \text{ nm}^{-1}$  this was not necessary because it does not influence the position of the LA phonons. This elastic contribution is  $Q$ -dependent and is probably due to residual

disorder in the polycrystal. *b*, Inelastic X-ray scattering spectra of liquid  $H_2O$  at  $4^\circ \text{C}$ , taken at the  $Q$  values indicated. The data are normalized to the intensity of the central peak maximum, and the count rates were roughly 5, 4, 5 and 6  $\text{count s}^{-1}$ , at  $Q = 4, 6, 8$  and  $10 \text{ nm}^{-1}$ , respectively. The integration time for each data point was 120 s. The data points ( $\circ$ ), shown with error bars, are superimposed on the fit (solid line). The curve under the central peaks is the experimental resolution function, modelled by a lorentzian with 3.2 meV FWHM, which is shown to emphasize the  $Q$ -dependent intensity on the sides of the central peaks. The fit was obtained by the convolution of the resolution function with a lorentzian for the central peak, and a damped harmonic oscillator lineshape<sup>2,12</sup> for the two side peaks. In the fits shown in *a* and *b*, detailed balance has always been imposed<sup>2</sup>.



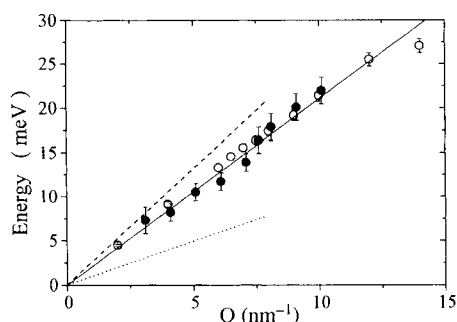


FIG. 2 Dispersion relation in solid (○) and liquid (●) water as obtained from the fits shown in Fig. 1. The solid line indicates the best linear fit over the range  $Q = 4\text{--}10\text{ nm}^{-1}$ , and the slope corresponds to a speed of sound of  $3,200 \pm 100\text{ m s}^{-1}$ . The dashed line has a slope of  $4,000\text{ m s}^{-1}$  and corresponds to the speed of sound of ice Ih in the ultrasonic region. The dotted line has a slope of  $1,500\text{ m s}^{-1}$ , and corresponds to the speed of sound of liquid water in the hydrodynamic regime.

maximum of the lorentzian representing the LA phonon. In the liquid, it is the  $\Omega(Q)$  parameter of the damped harmonic oscillator model as defined in ref. 12. The ice Ih data are consistent with previous neutron measurements in  $\text{D}_2\text{O}$  made in higher Brillouin zones<sup>13</sup>.

The two dispersion relations in Fig. 2 are strikingly similar. They are in fact identical within the reported error bars. This result demonstrates that in the  $Q$  region investigated, collective longitudinal excitations propagate with the same velocity in the liquid and solid phase of  $\text{H}_2\text{O}$ . This velocity,  $\sim 3,200\text{ m s}^{-1}$  in the  $Q$  range  $4\text{--}14\text{ nm}^{-1}$ , is lower than the value of  $4,000\text{ m s}^{-1}$  observed in polycrystalline ice Ih at ultrasonic frequencies, and more than twice the value of  $\sim 1,500\text{ m s}^{-1}$  found in water in the hydrodynamic limit. This is emphasized by the dashed and dotted lines in Fig. 2, which are extrapolations of the dispersions measured respectively in ice Ih and liquid water at  $Q$  smaller than  $10^{-2}\text{ nm}^{-1}$ . The origin of this strong frequency dependence of the sound velocity in liquid water, a phenomenon known as fast sound, is still highly debated<sup>14–17</sup>.

A theoretical model consistent with all the experimental findings so far associated with collective dynamics in water explains the fast sound as an upward bending of the dispersion relation. This bending is anomalously high in liquid water, where the interplay between the electrostatic and Lennard–Jones-like contributions to the intermolecular potential produces a large infinite-frequency sound velocity<sup>16</sup>. This quantity characterizes the solid-like elastic response of the liquid, independent of viscosity-related processes. The equivalence of the sound velocity in liquid and solid water indicates, therefore, that this infinite-frequency limit has already been reached in the liquid at energies below  $\varepsilon \approx 5\text{ meV}$ , and it implies that in water the relaxation times of all viscous dissipation processes are longer than  $\hbar/\varepsilon \approx 100\text{ fs}$ .

Our measurements indicate that, in water, there exists a space–time domain in the mesoscopic region where the interparticle correlations are very similar in the liquid and solid phases. This result, so far, is specific to water, and it will be of great interest to investigate the behaviour of other liquids. Finally, our results put further constraints on theories that try to account for the phenomenon of fast sound. □

Received 6 September; accepted 22 December 1995.

- Teixeira, J., Bellissant-Funel, M. C., Chen, S. H. & Dörner, B. *Phys. Rev. Lett.* **54**, 2681–2683 (1985).
- Sette, F. et al. *Phys. Rev. Lett.* **75**, 850–853 (1995).
- Van Hove, L. *Phys. Rev.* **95**, 249–262 (1954).
- Bottom, V. E. A. *Acad. Brasileira Ciências* **37**, 407–413 (1965).
- Maier-Leibnitz, H. *Nucl. Instrum. Meth.* **8**, 61–67 (1966).
- Graeff, W. & Materlik, G. *Nucl. Instrum. Meth.* **195**, 97–103 (1982).
- Dörner, B., Burkel, E. & Peisl, J. *Nucl. Instrum. Meth.* **A246**, 450–451 (1986).
- Masciovecchio, C. et al. *Nucl. Instrum. Meth.* (in the press).

- Verbeni, R. et al. *J. Synch. Radiation* (in the press).
- Bosi, P., Tubino, R. & Zerbi, G. *J. chem. Phys.* **59**, 4578–4586 (1973).
- Hansen, J. P. & McDonald, I. R. *Theory of Simple Liquids*, 229–236 (Academic, London, 1976).
- Fak, B. & Dörner, B. *Institute Laue Langevin Tech. Rep. No. 92FA008S* (1992).
- Renker, B. *Phys. Lett.* **A30**, 493–494 (1969).
- Rahman, A. & Stillinger, F. H. *Phys. Rev.* **A10**, 368–378 (1974).
- Impey, R. W., Madden, P. A. & McDonald, I. R. *Molec. Phys.* **46**, 513–539 (1982).
- Balucani, U., Ruocco, G., Torcini, A. & Vallauri, R. *Phys. Rev.* **E47**, 1677–1684 (1993).
- Sciortino, F. & Sastry, S. *J. chem. Phys.* **100**, 3881–3893 (1994).

ACKNOWLEDGEMENTS. We thank O. Consorte, W. Galli, B. Gorges, K. Martell and J. F. Ribois for their technical assistance, and the ESRF staff for help in commissioning the inelastic X-ray scattering beamline.

## A microscopic model for surface-induced diamond-to-graphite transitions

Alessandro De Vita\*, Giulia Galli\*, Andrew Canning† & Roberto Car\*

\* Institut Romand de Recherche Numérique en Physique des Matériaux, IN-Ecublens, CH-1015 Lausanne, Switzerland

† CRAY Research, Switzerland

GRAPHITIZATION of diamond at ambient pressure was first observed in the 1920s<sup>1,2</sup>, but the mechanisms responsible for this transformation and, in particular, those underlying the nucleation and growth of graphite in diamond, remain controversial<sup>3–5</sup>. In addition to their fundamental interest, these processes have technological relevance—for example, for the growth by chemical vapour deposition<sup>6</sup> of diamond-like films, which sometimes include graphitic islands<sup>7</sup>. Here we report the results of first-principles molecular dynamics simulations of a surface-induced diamond-to-graphite transition, which provide a microscopic model for the early stages of the graphitization process. We find that a well defined diamond/graphite interface forms during the transition; the electronic properties of the atoms at this interface suggest that they are highly chemically active sites. In addition to its relevance to graphite inclusion in diamond films, our model should yield insight into the process of selective etching in vapour-deposited carbon films, and possibly also into diamond nucleation.

The first theoretical model for the diamond-to-graphite transition is due to Nath<sup>3</sup>, who proposed a direct transformation path between the two phases. According to his model, at high temperature, the puckered (111) diamond planes flatten, and the resulting hexagonal planes (in a rhombohedral A,B,C stacking) increase their relative spacing and then glide to form hexagonal (A,B stacked) graphite.

Later experimental work<sup>4,5</sup> clarified the surface-induced nature of the diamond-to-graphite transition and confirmed previous observations<sup>8,9</sup> on the presence of diamond(111) planes oriented parallel to graphite(0001). Davies and Evans<sup>5</sup> dismissed as very unlikely the possibility of a direct martensitic transformation from diamond to graphite. In agreement with earlier suggestions<sup>10–12</sup>, they instead proposed<sup>4,5</sup> a two-stage transition mechanism in which carbon atoms evaporate from the diamond surface to form disordered aggregates which then condense into graphitic micrograins. In this model, graphite crystallites subsequently form by a grain growth process.

We have addressed the issue of the microscopic mechanism underlying the surface-induced graphitization of diamond by performing first-principles molecular dynamics simulations of bulk diamond, terminated by a (111) surface, as a function of temperature. In first-principles simulations, the forces acting on atoms which are needed for solving the newtonian equations of motion are obtained directly from fundamental quantum-

mechanical theory<sup>13</sup>. We performed all our computations using periodic boundary conditions on repeated diamond slabs containing twelve (111) atomic layers of carbon. The slabs are parallel to the  $x$ - $y$  plane of the orthorhombic unit cell, and are separated by  $\sim 9$  Å in the  $z$  direction. In one set of computations the repeated cell contains 192 carbon atoms, that is 16 carbon atoms per layer. The  $x$ ,  $y$  and  $z$  cell dimensions are 8.74 Å, 10.10 Å and 21.17 Å respectively, the  $x$  and  $y$  dimensions being oriented along the [110] crystallographic axes of cubic diamond (for which the experimental lattice constant of 3.57 Å is used). The slab is initially terminated on both sides by a ( $2 \times 1$ ) reconstructed surface, that is, the stable surface geometry at 0 K (ref. 14). (Fig. 1a). This reconstruction is observed experimentally after dehydrogenation and is believed to be the stable geometry of the (111) surface at temperatures relevant to graphitization. Another set of computations was carried out on the same system terminating one of the slab surfaces by hydrogen atoms. A third set of calculations was performed on a system consisting of 288 carbon atoms, whose cell geometry is obtained from the one described above by increasing the cell size in the  $x$  direction by a factor of 1.5. First-principles simulations with such large cells and for a total simulation time of about 15 ps were made possible using a massively parallel computer and a novel optimized code<sup>15</sup>.

In the first part of our study, the 192-atom slab was heated up to 2,500 K by carrying out constant-temperature simulations<sup>16,17</sup>. At  $\sim 2,500$  K, a surface-induced graphitization of the diamond slab was observed. Some of the long<sup>14</sup> bonds between the first and the second double layers break, and the first double layer flattens to form a dome-shaped strained graphitic seed (Fig. 1b). Once the seed is nucleated, fast graphitization of the whole slab takes place. Graphitization proceeds by correlated breaking of nearest-neighbour bonds oriented along the  $z$  direction. This implies a paired  $sp^3$  to  $sp^2$  hybridization change of nearest-neighbour atoms lying in the  $x$ - $y$  plane, presumably favoured by the energy gained in the formation of  $\pi$ -bonds in the plane. Hence the graphite phase penetrates into diamond before full graphite planes are formed (Fig. 1c), and a coherent diamond/graphite interface, not aligned with the surface, is formed during the transition process. The graphite stacking obtained at the end of the transition (Fig. 1d) is (A,B,C), that is, the same as that of the starting diamond slab. The size of our simulation prevents the study of a possible stacking change after the phase transition.

To substantiate our findings, we repeated some of our calculations with the slab terminated by a reconstructed surface on one side, and by hydrogen atoms on the other side. The hydrogen-terminated side (H plus one unreconstructed double layer of C) was kept fixed during the simulations, to mimic a rigid bulk attachment. We found that the graphitization process occurred in the same way as in the system with the slab terminated on both sides by a surface. Our results indicate that graphitic microcrystals can form at the diamond surface by a direct transformation, in spite of the dissimilarities between the structural and electronic properties of diamond and graphite. Formation of graphitic microcrystals by evaporation and condensation would probably require a larger activation energy. We note that in a macroscopic sample, such microcrystals could not grow indefinitely because of the large strain energy associated with the diamond/graphite interface. It is conceivable that stress can be relieved by, for example, formation of defects breaking the coherence of the interface. This would lead to the formation of graphitic micrograins, with a substantial roughening of the graphitized surface as observed experimentally. Subsequent grain growth could then occur in a way similar to that indicated by Davies and Evans<sup>5</sup>. These speculations could not be tested in our simulation, which was limited by time and size to the early stage of the graphitization process.

In recent years, the debate over the diamond-to-graphite transformation has gained technological relevance in the context of low-pressure growth of diamond films by chemical vapour deposition (CVD) techniques<sup>6</sup>. When grown under high-tempera-

ture conditions or from carbon-rich hydrocarbon plasma, diamond films are found to include graphitic islands<sup>7</sup>. Some experimental observations<sup>7</sup> indicate that these islands coherently match the diamond phase by forming interfaces characterized by diamond(111) planes parallel to graphite(0001) planes. The same epitaxial relationship between diamond and graphite has also been observed in diamond deposited by CVD on graphite substrates<sup>18</sup>, where diamond nucleation could proceed by hydrogenation of graphitic precursors<sup>19</sup>. The coexistence of diamond and graphite, and the growth of one phase from the other, has so far only been studied with empirical models. In particular, Lambrecht *et al.*<sup>19</sup> proposed a model for diamond growth at both non-hydrogenated and hydrogenated prism planes of graphite formations. Klein<sup>20</sup> used a tight-binding picture to describe the electronic structure of graphitic polymer strips with edge states, and of diamond-graphite hybrids<sup>21</sup> which could be present in regions of coexistence of the diamond and graphite phases. Tight-binding models have also been used to study the possible competition between reconstruction and graphitization at clean<sup>22,23</sup> and stepped surfaces<sup>24</sup>.

In order to investigate how diamond and graphite can coexist, we performed a series of constant-temperature simulations which were started from atomic configurations containing a frozen seed of the diamond-to-graphite transition found in our previous computer simulations. We first carried out simulations with the 192-atom slab at 1,000 K. These calculations clearly showed the formation of an interface, between the diamond and graphite phases, which was stable over the time span of our simulation. With the aim of reducing finite-size effects, we carried out additional calculations on a larger 288-atom slab, starting from the diamond/graphite interface geometry obtained with the smaller slab. These simulations showed that the interface was stable in the temperature range 0–1,000 K. The interface, shown in Fig. 2, is characterized by the alternation of three- and four-fold coordinated atoms along the [110] direction. The merging of the two phases into each other is sharp, occurring over one bond length.

The electronic properties of the interface were investigated by computing the local density of electronic states (LDOS). One of the prominent features of the LDOS is a sharp maximum close to the Fermi level in the region of the interface at the edge of the graphite planes. This peak arises from electronic states with a predominantly  $p_z$  character, localized at the diamond/graphite interface. In particular, the lowest unoccupied states are localized at the insertion of the graphite plane into the first and second surface layers of the diamond (see Fig. 3). Thus, we expect these surface atoms to act as chemically active sites, for hydrogen chemisorption for example. This is consistent with the results of band structure calculations of hydrogenated graphite<sup>25</sup>. Similarly, in CVD-grown diamond films, dangling bonds should be present at the surface edge of a graphite crystallite growing into diamond. As originally reported by Angus *et al.*<sup>26</sup>, graphite formations can be preferentially removed from diamond like films by etching with atomic hydrogen. Our prediction of preferential hydrogenation sites at the coherent insertion of the (stable) graphitic phase into the (metastable) diamond phase clearly substantiates the role of

FIG. 2 Diamond/graphite interface obtained from a simulation on the 288-atom system. The system was structurally relaxed (at 0 K) starting from an interface configuration obtained from a constant-temperature simulation at  $\sim 1,000$  K on the 192-atom slab (see text). The system was subsequently heated up to 300 K. This figure represents the equilibrium structure obtained after  $\sim 1$  ps of simulation at 300 K. The ring networks of the curved graphite sheets are coherently connected to the buckled diamond (111) double layers. Notice the alternation of three- and four-fold coordinated atoms along the [110] rows at the convex corner of the phase boundary (centre of the slab, colour scheme as defined in Fig. 1 legend). This interface reconstruction maximizes the number of closed six-fold rings of  $sp^2$ -bonded atoms on the edge of the graphite sheets at the phase boundary.

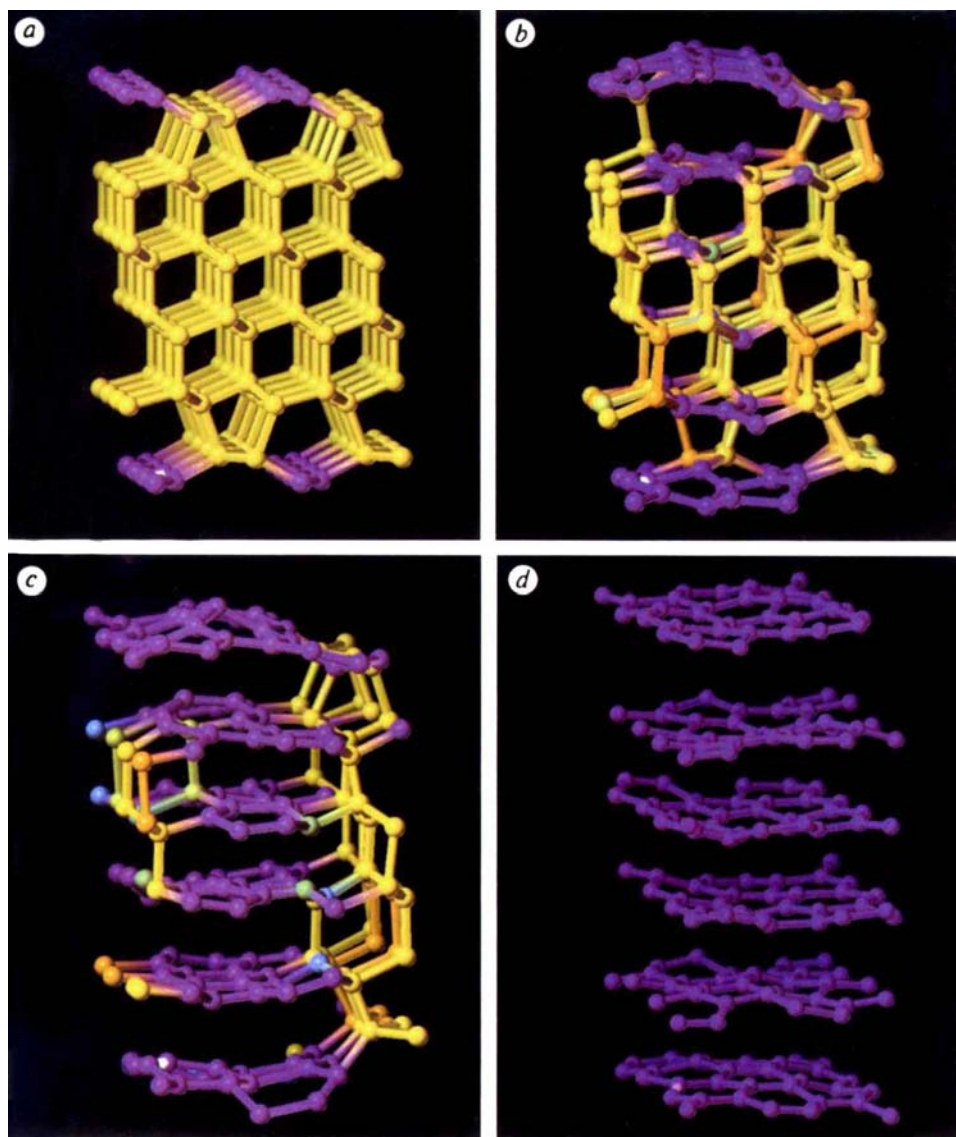
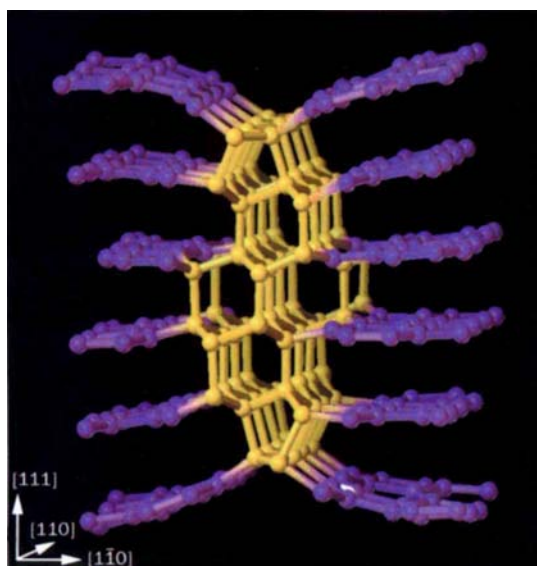


FIG. 1 Surface-induced graphitization process of the 192-atom diamond slab, at  $\geq 2,500$  K. (Yellow spheres, diamond; purple spheres, graphite.) The system is represented by a periodically repeated orthorhombic cell containing a slab of six (111) double layers of C atoms, separated by a vacuum region  $\sim 9$  Å thick. The three-fold and four-fold coordinated atoms belonging to one cell are represented by purple and yellow spheres. Orange, green and blue spheres represent atoms of progressively lower coordination, intermediate between those of diamond and graphite. *a*, The original  $(2 \times 1)$  reconstructed slab at 0 K. *b*, Initial stage of graphitization. The process starts at the surface some bonds between the first and the second double layer are broken, and a dome-shaped strained graphitic seed is formed (at time  $t \approx 0.1$  ps). *c*, Penetration of the graphite phase into the diamond slab ( $t \approx 0.3$  ps). In the bond network of the diamond subsurface region every z-oriented bond connects an atom to a neighbouring atom in the (111) double layer directly underneath, whose three remaining neighbours belong to the same double layer. These three remaining neighbours, in turn, participate in three more z-oriented bonds located further into the slab. Thus correlated breaking of z-oriented bonds close to each other induces penetration of the graphite phase into diamond before full graphite planes are formed. *d*, End of the phase transition ( $t \approx 0.5$  ps), giving defect-free (A,B,C stacked) graphite planes. The graphitic structure is under slight tensile stress ( $\sim 2\%$ ) in the x-y plane because of the constraint on the planar lattice constant. Such a constraint would be also present in experiments, during the initial stage of graphitization, owing to the underlying diamond substrate. A diamond-to-graphite transition was also obtained at lower temperature, by quenching the

system to  $\sim 1,900$  K after the nucleation of the initial graphitic seed (*b*). The reaction then proceeded in the same way as at  $\geq 2,500$  K. The interaction between ionic cores and valence electrons is described by a fully non-local pseudopotential<sup>27</sup> with s-only non-locality. The single particle orbitals and charge density are expanded in plane waves with kinetic energy cutoffs of 33 and 99 Ry, respectively. Repeating the calculation with a pseudopotential optimised for basis set convergence<sup>28</sup>, and using a cutoff of 40 Ry for the orbital expansions, graphitization proceeded in the same way as illustrated in the figure. Only Bloch functions at the  $\Gamma$ -point of the cell Brillouin zone are included in the calculations. The thermodynamic driving force of the transition is given by the free-energy difference  $\Delta F = F(D) - F(G)$  (where  $F(D)$  and  $F(G)$  are the free energies of diamond and graphite, respectively), which is dominated by the vibrational properties of the two phases. A simple estimate based on the harmonic formula and on the known phonon frequencies of diamond and graphite gives a vibrational  $\Delta F_v \approx 0.18$  eV per atom at 3,000 K. On this scale, diamond and graphite are almost degenerate in energy at 0 K. Using our standard choice of parameters and suitable 192-atom bulk cells having the same x and y dimensions as our simulation cell, we obtained  $\Delta E = -0.04$  eV per atom for the energy difference at 0 K between diamond and rhombohedral (strained) graphite. A well converged total energy difference between the two phases in equilibrium at 0 K is<sup>29</sup>  $-0.01$  eV per atom, neglecting zero-point motion ( $\sim 0.015$  eV per atom). For comparison, the experimental energy difference between diamond and hexagonal graphite is<sup>30</sup> 0.025 eV per atom.





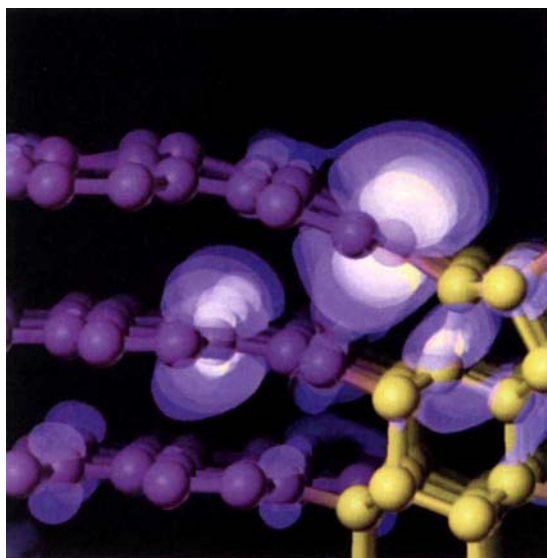


FIG. 3 Isovalue surfaces of electron density corresponding to an electronic state localized at the interface between the diamond and graphite phases. The figure shows a section of the system of Fig. 2, corresponding to the insertion into diamond of the first few graphite sheets. (Two isovalue surfaces are shown, in yellow and blue, corresponding to high and low values of the electronic density, respectively). The electronic state illustrated is the lowest unoccupied one, its energy eigenvalue being only  $\sim 0.02$  eV above the Fermi energy in our finite-size calculation. Although the orbital is partially spread into the sub-surface region of both graphite and diamond in the vicinity of the interface, its prominent feature consists of the large density lobes located at the insertion into diamond of the first graphitic layer. The density lobes have a clear  $p$  character, and point into the vacuum region. Where this state is found, the surface C atoms are expected to act as chemically active sites (for hydrogen chemisorption, for example). The presence of this state (and other localized electronic states of similar energy) at the diamond/graphite surface phase boundary could play an important role in atomic-H etching of any graphitic nuclei which may form in diamond-like films.

hydrogen in inhibiting graphite growth. We also note that the electronic properties of diamond-graphite hybrids, as derived from empirical models<sup>20,21</sup>, are consistent with our results. Indeed, Balaban and co-workers<sup>21</sup> suggested the existence of localized electronic states at a diamond/graphite interface resulting from the coupling between the  $\pi$ -like graphitic states and the  $\sigma$ -like diamond states. They also suggested<sup>21</sup> that semimetallic or metallic conduction might be observed along the diamond/graphite interface, which is consistent with our findings.  $\square$

Received 12 June 1995; accepted 3 January 1996.

1. Friedel, G. & Ribaud, G. *C. r. hebd. Séanc. Acad. Sci., Paris* **178**, 1126–1129 (1924).
2. Libeau, P. & Picon, M. *C. r. hebd. Séanc. Acad. Sci., Paris* **179**, 1059–1061 (1924).
3. Nath, N. S. N. *Proc. Indian Acad. Sci.* **A2**, 143–152 (1935).
4. Evans, T. & James, P. F. *Proc. R. Soc. Lond.* **A277**, 260–269 (1964).
5. Davies, G. & Evans, T. *Proc. R. Soc. Lond.* **A328**, 413–427 (1972).
6. Spear, K. E. & Frenklach, M. *Pure appl. Chem.* **66**, 1773–1782 (1994).
7. Zhu, W., Rendall, C. A., Badzian, A. R. & Messier, R. J. *Vac. Sci. Technol.* **A7**, 2315–2324 (1989).
8. Grenville-Wells, H. J. *Mineralog. Mag.* **29**, 803–816 (1952).
9. Seal, M. *Proc. 4th Int. Conf. on Electron Microscopy* (ed. Ross, R.) 455–459 (Royal Microscopical Soc., London, 1958).
10. Bovenkerk, H. P., Bundy, F. P., Hall, H. T., Strong, H. M. & Wentorf, R. H. *Nature* **184**, 1094–1098 (1959).
11. Rodewald, H. J. *Helv. chim. Acta* **43**, 1657–1666 (1960).
12. Bundy, F. P., Hall, H. T., Strong, H. M. & Wentorf, R. H. *J. chem. Phys.* **35**, 383–391 (1961).
13. Car, R. & Parrinello, M. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
14. Iarlori, S., Galli, G., Gygi, F., Parrinello, M. & Tosatti, E. *Phys. Rev. Lett.* **69**, 2947–2950 (1992).
15. De Vita, A. et al. *EPFL Superconducting J.* **6**, 22–27 (1994).
16. Nosé, S. *Molec. Phys.* **52**, 255–262 (1984).
17. Hoover, W. *Phys. Rev.* **A31**, 1695–1697 (1985).
18. Li, Z. et al. *J. appl. Phys.* **73**, 711–715 (1993).
19. Lambrecht, W. R. L. et al. *Nature* **364**, 607–609 (1993).
20. Klein, D. J. *Chem. Phys. Lett.* **217**, 261–265 (1993).
21. Balaban, A. T., Klein, D. J. & Folden, C. A. *Chem. Phys. Lett.* **217**, 266–270 (1993).
22. Jungnickel, G. et al. *Bull. Am. phys. Soc.* **40**, 772 (1995).
23. Jungnickel, G. et al. *Mater. Res. Symp. Proc.* **383**, 349–360 (1995).
24. Davidson, B. & Pickett, W. *Phys. Rev.* **B49**, 11253–11256 (1994).
25. Mehndru, S. P., Anderson, A. B. & Angus, J. C. *J. phys. Chem.* **96**, 10978–10982 (1992).
26. Angus, J. C., Will, H. A. & Stanko, W. S. *J. appl. Phys.* **39**, 2915–2922 (1968).
27. Bachelet, G. B., Hamann, D. & Schlüter, M. *Phys. Rev.* **B26**, 4199–4228 (1982).
28. Troullier, N. & Martins, J. L. *Phys. Rev.* **B43**, 1993–2006 (1991).
29. Fahy, S., Louie, S. & Cohen, M. L. *Phys. Rev.* **B34**, 1191–1199 (1986).
30. Weast, R. C. (ed.) *CRC Handbook of Chemistry and Physics* 75th edn D-58 (CRC, Boca Raton, Florida 1988).

ACKNOWLEDGEMENTS. This work was partially supported through the Parallel Application Technology Program (PATP) between the EPFL and Cray Research, Inc. and by the Swiss NSF.

## Satellite confirmation of the dominance of chlorofluorocarbons in the global stratospheric chlorine budget

James M. Russell III\*, Mingzhao Luo†, Ralph J. Cicerone† & Lance E. Deaver\*

\* NASA, Langley Research Center, Hampton, Virginia, 23681-0001, USA  
† University of California, Irvine, California 92717-3100, USA

OBSERVED increases in concentrations of chlorine in the stratosphere<sup>1–7</sup> have been widely implicated in the depletion of lower-stratospheric ozone over the past two decades<sup>8–14</sup>. The present concentration of stratospheric chlorine is more than five times that expected from known natural ‘background’ emissions from the oceans and biomass burning<sup>15–18</sup>, and the balance has been estimated to be dominantly anthropogenic in origin, primarily due to the breakdown products of chlorofluorocarbons (CFCs)<sup>19,20</sup>. But despite the wealth of scientific data linking chlorofluorocarbon emissions to the observed chlorine increases,

the political sensitivity of the ozone-depletion issue has generated a re-examination of the evidence<sup>21,22</sup>. Here we report a four-year global time series of satellite observations of hydrogen chloride (HCl) and hydrogen fluoride (HF) in the stratosphere, which shows conclusively that chlorofluorocarbon releases—rather than other anthropogenic or natural emissions—are responsible for the recent global increases in stratospheric chlorine concentrations. Moreover, all but a few per cent of observed stratospheric chlorine amounts can be accounted for by known natural and anthropogenic tropospheric emissions. Altogether, these results implicate the chlorofluorocarbons beyond reasonable doubt as dominating ozone depletion in the lower stratosphere.

The Halogen Occultation Experiment (HALOE), described in detail elsewhere<sup>23</sup>, has been operating from the Upper Atmosphere Research Satellite platform almost continuously since it was first turned on in orbit on 11 October 1991. It uses the technique of solar occultation to provide vertical profiles through the global stratosphere of HCl, HF, O<sub>3</sub>, CH<sub>4</sub>, H<sub>2</sub>O, NO, NO<sub>2</sub>, aerosol extinction and temperature. All HALOE data have undergone an extensive validation process that included a number of internal data consistency studies, comparisons with model simulations and detailed comparisons with remote observations made from the ground, balloon, aircraft and satellite platforms. Results for HF (ref. 24) show the on-orbit data precision to be  $\sim 0.04$ – $0.06$  parts per 10<sup>9</sup> by volume (p.p.b.v.) throughout the stratosphere and

lower mesosphere. The estimated accuracy is 14–27% depending on altitude. These estimates appear to be quite conservative because comparisons with a set of nine HF profiles measured in balloon underflights show mean differences of only <7% from 5 to 50 mbar which is well within the error bar overlap of the two measurement sets. In general the HF results are very robust. HCl studies<sup>25</sup> show an on-orbit precision of 0.1–0.2 p.p.b.v. and the estimated accuracy ranges from 12% to 24% depending on altitude. The mean difference between HALOE HCl and results from a set of 14 balloon underflights varies from 8% to 16% over the 5–70 mbar range; HALOE values are lower than the balloon

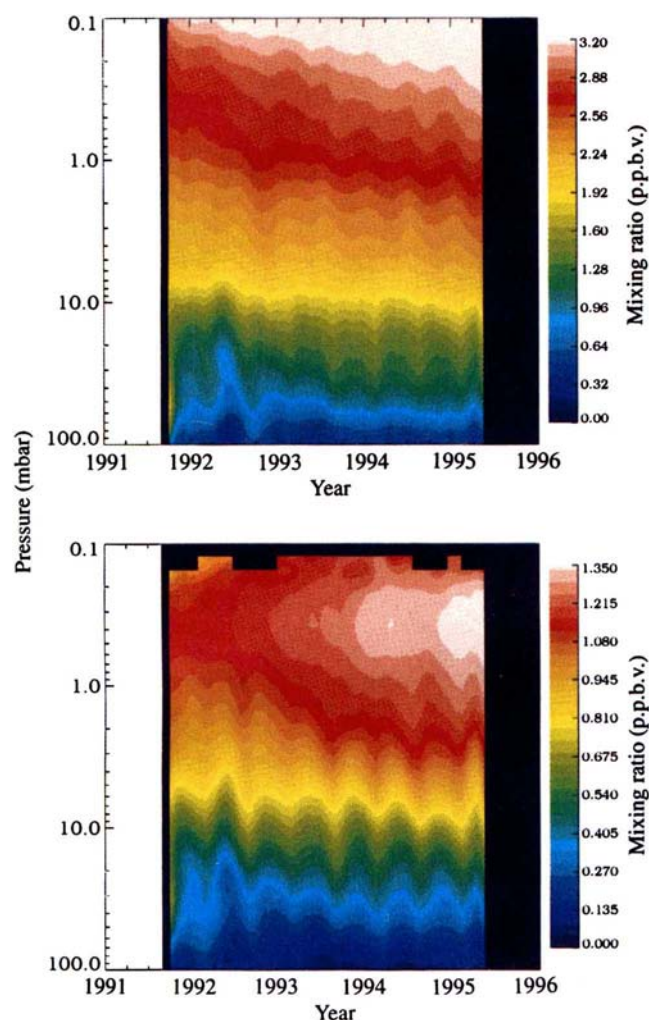


FIG. 1 HALOE HCl (top) and HF (bottom) time series over the range from 100 mbar (roughly the tropopause) to 0.1 mbar (~65 km) using data collected at all  $\beta$  angles. The data have been globally and zonally averaged. On any given day, all sunrise and sunset data are zonally averaged separately and then a combined average is obtained without regard to latitude location of the sunrise or sunset measurements. Temporal increases in HCl and HF are clear at the highest altitudes near the top of each panel but with careful inspection they can be seen in the mid-stratosphere as well. At 0.3 mbar, for example, HCl starts out at ~2.7 p.p.b.v. and at the end of the record it reaches ~3.1 p.p.b.v. HF varies similarly from 1.1 p.p.b.v. to ~1.4 p.p.b.v. over the period. We note that the half-yearly variations in both HCl and HF that are most apparent near the 10 mbar level are due to a sampling artefact of the occultation geometry, causing certain latitudes to be preferentially sampled at a given time of year. The high HCl and HF in the lower stratosphere at the start of the data record as due to oversampling of high latitudes during this period. The 15 mbar level HCl varies with latitude from ~1.6 p.p.b.v. in the tropics to 2.1 p.p.b.v. at high latitudes. HF varies from 0.3 p.p.b.v. to 1 p.p.b.v. over this same range. Note that around the stratopause, latitudinal changes in both HCl and HF are small.

data. A HCl data artefact appears as a dependence of HCl on  $\beta$ , the angle between the orbit plane and the Earth–Sun line. HALOE HCl varies in a systematic way in the upper stratosphere by  $\pm 9\%$  over a  $\beta$  angle cycle. The instrumental cause of this artefact has now been identified and it will be removed in future data processing. As we shall show later, the  $\beta$  angle variation has negligible effect on HCl annual increase rates determined from HALOE data. The important parameter for time trend comparisons is data precision, which is high for both HF and HCl.

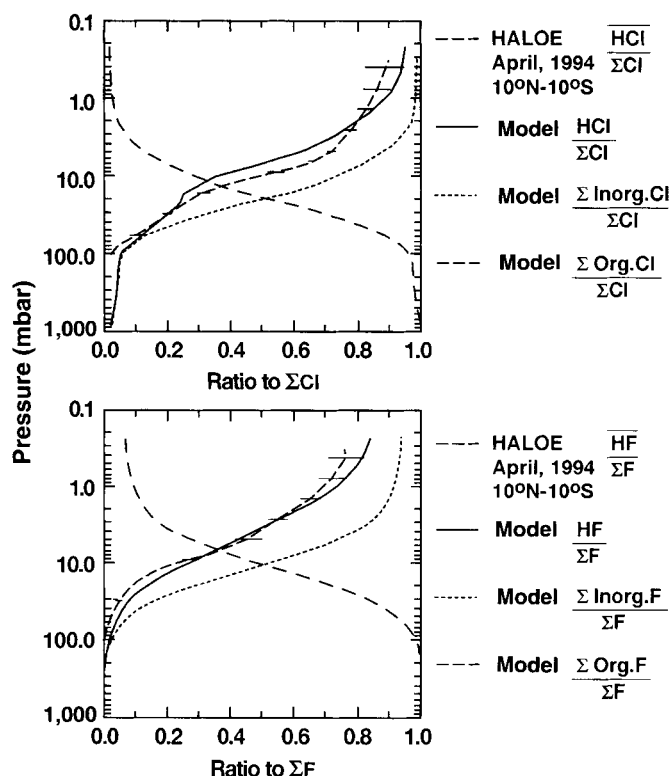


FIG. 2 A steady-state model simulation at the equator for 1990 halocarbon tropospheric boundary conditions showing altitude-dependent model ratios of HCl, HF, inorganic Cl and F, and organic Cl and F to model  $\Sigma\text{Cl}$  and  $\Sigma\text{F}$  obtained by using the NCAR two-dimensional chemistry-transport model<sup>32</sup>. The HALOE HCl and HF over model  $\Sigma\text{Cl}$  and  $\Sigma\text{F}$  ratios are also shown. Note that near the Earth surface,  $\Sigma\text{Cl}$  and  $\Sigma\text{F}$  equal  $\text{Cl}_T$  and  $\text{F}_T$  respectively. As altitude increases, the organic compounds decompose owing to photolysis and chemistry forming inorganic compounds until in the upper stratosphere and higher they dominate  $\Sigma\text{Cl}$  and  $\Sigma\text{F}$ . At these altitudes  $\Sigma\text{Cl} = \text{HCl} + \text{Cl} + \text{ClO} + \text{ClONO}_2 + \text{HOCl} + \dots$  and  $\Sigma\text{F} = \text{HF} + 2\text{CF}_2\text{O} + \text{CClFO} + 2\text{CHF}_2\text{Cl} + \dots$ . The HALOE data were collected in 1994 and averaged over the range  $10^\circ\text{S}$  to  $10^\circ\text{N}$ . The 1994 HALOE measurements are compared with 1990 model steady-state data to roughly take into account the lag time for chlorine and fluorine in the troposphere to reach the upper stratosphere and lower mesosphere. We note that the vertical profile shapes of HALOE data and model simulations are in good agreement. These results show that above the  $p = 1$  mbar level, nearly all the Cl and F that can be converted to the inorganic forms HCl and HF has been converted, as seen by the nearly constant HCl/ $\Sigma\text{Cl}$  and HF/ $\Sigma\text{F}$  ratios, which have values approaching 1.0. The cause of residual Cl and F (that is, ratios <1.0) above the  $p = 1$  mbar level is mainly HCFC-22, which has a stratospheric lifetime of over 200 yr (ref. 33). In this altitude range, HCFC-22 comprises only 1–2% of  $\Sigma\text{Cl}$  but 2–5% of  $\Sigma\text{F}$ . Because the HCl and HF mixing ratios are nearly constant with altitude for  $p < 1$  mbar, seasonal and dynamical effects are less important and to first order can be ignored in trend determinations. For these reasons we selected 55 km as the basis for the time series comparisons (for example,  $\text{Cl}_T(t)$  against  $\text{HCl}_T(t)$ ). At this altitude, HCl and HF levels are ~93% and ~81% of  $\Sigma\text{Cl}$  and  $\Sigma\text{F}$  levels respectively at the equator. These ratios increase to ~95% and ~88% at high latitudes. Other model results<sup>34</sup> give a HF to  $\Sigma$  (inorganic F) ratio (that is, HCHC-22 is neglected) of ~90%. Spacelab 3 ATMOS data for  $28^\circ\text{N}$  (ref. 5) show ratios at 55 km of  $95 \pm 5\%$  and  $79 \pm 14\%$  for HCl/ $\Sigma\text{Cl}$  and HF/ $\Sigma\text{F}$  respectively, in good agreement with model results.

Because stratospheric HF has no significant known natural sources<sup>17,26</sup> its observation in the stratosphere is especially important as positive proof of transport of CFCs from the troposphere to the stratosphere, their photolysis, and liberation of chlorine and fluorine. For comparison with our observations of HF and HCl in the stratosphere, we have calculated tropospheric mixing ratios of total organic chlorine ( $Cl_T$ ) and total organic fluorine ( $F_T$ ) on the basis of concentrations of key gases measured at over 20 stations worldwide<sup>27,28</sup>. We calculated  $Cl_T$  and  $F_T$  using the Cl or F atom-weighted sum for the years 1985–87, 1989, 1990 and 1992. We omitted  $CF_4$  and  $C_2F_6$  in the calculation because, owing to their extremely long lifetimes ( $>1,000$  yr; ref. 28), they do not contribute significantly to the stratospheric inorganic fluorine amount. We have not included the recently introduced CFC substitutes HCFC-142b ( $CH_3CClF_2$ ), HCFC-141b ( $CH_3CCl_2F$ ), and HFC-134a ( $CF_3CH_2F$ ) because their pre-1990 concentrations are very small compared with the 1992 CFC levels and in addition the available data on these compounds cover only the period 1991–94 (ref. 29). Also we ignore tropospheric HCl and HF from sources other than organic compounds, an assumption that will be seen to be valid in our results. All other important organic-chlorine and fluorine compounds with a lifetime  $>1$  year were included (see Fig. 3). After making appropriate calibration adjustments<sup>30</sup>, the yearly total organic  $Cl_T$  and  $F_T$  data between 1985 and 1992 were fitted with linear regression lines to show the kinds of increasing trend that have been monitored near the Earth surface. On average, the fits show tropospheric total organic chlorine and

fluorine growing at the rates of  $116 \pm 12$  parts per  $10^{12}$  by volume (p.p.t.v.) and  $88 \pm 8$  p.p.t.v. per year respectively for the period 1985–92. Recent analyses show that the growth rates of CFC-11 and CFC-12 have significantly decreased because of the phasing out of production of these compounds as a result of international protocol agreements<sup>2</sup>. The growth rates of other halocarbons, including H-1301, H-1211 and  $CH_3CCl_3$ , are also decreasing<sup>30,31</sup>. Stratospheric chlorine arising from CFC-11 and CFC-12 should reach a maximum before the turn of the century and then begin to decline, assuming that the tropospheric growth rates of these two species continue to decrease as projected industrial emissions decline<sup>27</sup>.

Our first objective is to compare HALOE observations of HCl and HF rates of increase (Fig. 1) with tropospheric rates irrespective of absolute values or lag times for tropospheric  $Cl_T$  and  $F_T$  to be transported to the stratosphere. Such a comparison directly addresses the question of whether CFCs drive the stratospheric chlorine input, dominating over any natural input. The best region for comparison with  $Cl_T$  and  $F_T$  trends is in the upper altitudes, where almost all of the organic chlorine and fluorine have been converted to the inorganic forms, HCl and HF. We determined the altitude range where this has occurred (about 55 km,  $\sim 0.5$  mbar), using the NCAR two-dimensional chemistry transport model<sup>32</sup> to examine the changes in the ratio of HCl to  $\Sigma Cl$  and HF to  $\Sigma F$  with altitude (Fig. 2). The terms  $\Sigma Cl$  and  $\Sigma F$  are the atom-weighted sum totals of organic plus inorganic chlorine and fluorine respectively, again ignoring tropospheric HCl and HF from sources other than organic compounds. The trend comparisons (Fig. 3) show HALOE HCl and HF 55 km increase rates of  $102 \pm 6$  p.p.t.v.  $yr^{-1}$  and  $72 \pm 2$  p.p.t.v.  $yr^{-1}$  respectively, compared with  $116 \pm 12$  p.p.t.v.  $yr^{-1}$  and  $88 \pm 8$  p.p.t.v.  $yr^{-1}$  for  $Cl_T(t)$  and  $F_T(t)$ , where  $t$  indicates time. When the model ratios for HCl to  $\Sigma Cl$  and HF to  $\Sigma F$  are used to obtain HALOE-derived  $\Sigma Cl(t)$  and  $\Sigma F(t)$ , the rates are  $108 \pm 7$  p.p.t.v.  $yr^{-1}$  and  $85 \pm 2$  p.p.t.v.  $yr^{-1}$  respectively. This close agreement (3–7%) clearly shows that CFCs strongly, if not completely, dominate chlorine inputs to the stratosphere.

We can strengthen our argument by examining the HALOE derived chlorine amount at 55 km versus tropospheric organic chlorine,  $Cl_T$ . To do this we must take into account the time lag required for tropospheric gases to be transported to 55 km. We determined this by calculating the time separation between given

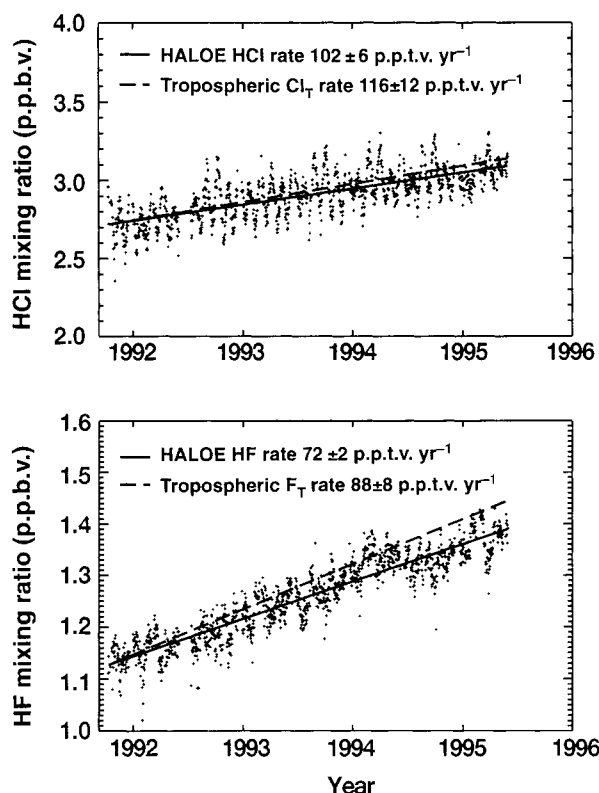


FIG. 3 HALOE global average HCl and HF time series at 55 km (solid lines) compared with tropospheric  $Cl_T$  and  $F_T$  time series (dashed lines) shifted in time using tropospheric starting values normalized to HALOE values taken at the beginning of the HALOE data record.  $Cl_T = 3CFCl_3 + 2CF_2Cl_2 + 3C_2Cl_3 + 2C_2F_2Cl_2 + C_2F_5Cl + CH_3Cl + 4CCl_4 + 3CH_3CCl_3 + CF_3BrCl + CHF_2Cl$ , and  $F_T = CFCl_3 + 2CF_2Cl_2 + 3C_2F_3Cl_3 + 4C_2F_4Cl_2 + 5C_2F_5Cl + 2CHF_2Cl + 2CF_2BrCl + 3CF_3Br$ . Data collected at all  $\beta$  angles were used. HALOE-derived  $\Sigma Cl(t)$  and  $\Sigma F(t)$  were obtained by linearly interpolating the model HCl to  $\Sigma Cl$  and HF to  $\Sigma F$  ratios with latitude and then applying these ratios to each HCl and HF data point. The effect of  $\beta$  angle dependence in the HCl data was evaluated by fitted only high- $\beta$  ( $> 45^\circ$ ) and low- $\beta$  ( $< 45^\circ$ ) data. These fits gave negligible differences for increase rates.

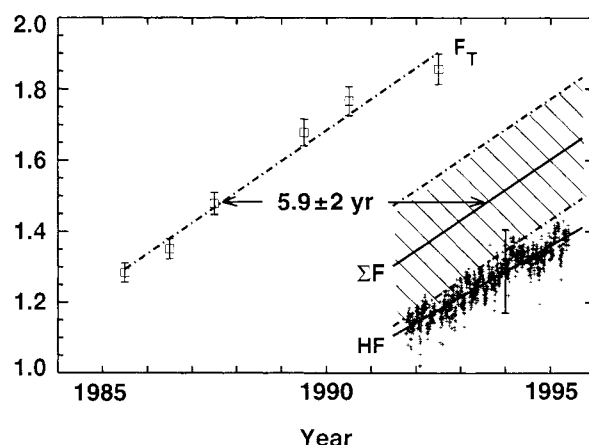


FIG. 4 Time series of HALOE HF at 55 km, HALOE-derived total organic plus inorganic fluorine ( $\Sigma F$ ), and tropospheric organic fluorine ( $F_T$ ). The hatched area represents the range of HALOE-derived  $\Sigma F$  when uncertainties associated with the straight-line fit, HALOE measurements and the model factor for HF/ $\Sigma F$  are considered. The vertical bar on the HF plot is the HALOE systematic error. The fluorine lag time of  $5.9 \pm 2$  yr is in good agreement with results inferred from ATMOS HF and  $CO_2$  (ref. 35) and the  $5.6 \pm 1.1$  yr deduced from balloon  $CO_2$  cryogenic samples<sup>36,37</sup> for transport to the middle stratosphere.



values in the HALOE-derived  $\Sigma F(t)$  time series and the same values in tropospheric  $F_T(t)$ , taking into account accuracies for both data sets (Fig. 4). We obtained a value of  $5.9 \pm 2$  yr. We chose the HF-derived lag time to do the chlorine budget calculation because it is an independently derived parameter and it is the proxy for the main chlorine carrier to the stratosphere, that is, the CFCs. Also, the HF measurement by HALOE is robust, it agrees very well with correlative underflight data, and it has no known artefacts to hinder the interpretation. We determined a 1 June 1995 expected total chlorine value at 55 km due to tropospheric emissions of  $3.44 \pm 0.23$  p.p.b.v.. This compares with HALOE-derived total chlorine,  $\Sigma Cl$ , of  $3.3 \pm 0.33$  p.p.b.v. using data for all  $\beta$  angles,  $3.44 \pm 0.34$  p.p.b.v. for  $\beta > 45^\circ$ , and  $3.22 \pm 0.33$  p.p.b.v. for  $\beta < 45^\circ$ . The HALOE mean  $\Sigma Cl$  value of 3.3 p.p.b.v. is less than the tropospheric value by only 6%. Also, the HALOE-derived  $\Sigma Cl$  trend and amount will increase slightly in the next round of data reprocessing for two reasons. First, the annual  $CO_2$  increase of  $\sim 0.5\%$  yr $^{-1}$  was not taken into account in the pressure registration algorithm used for inversion of the transmittance

data. Second, the  $CO_2$  band intensities used in the pressure registration algorithm will increase slightly owing to a spectroscopic reanalysis, which is nearing completion. The net result is that HALOE-derived  $\Sigma Cl$  at 55 km will agree with amounts predicted from tropospheric total organic chlorine to within a few per cent. This close agreement between the two data sets means that there are no important missing links in our understanding of the chlorine budget. Note that we ignored tropospheric HCl and HF from sources other than organic compounds in this analysis. If direct emissions of these compounds, surface or volcanic, were influencing the stratosphere, this excellent mass balance would not have been obtained. We further point out that the HALOE-derived mean total chlorine is  $\sim 2.7$  p.p.b.v. higher than the 0.6 p.p.b.v. level that it would reach if only the main natural source,  $CH_3Cl$ , were present. When these facts are considered together with observations showing that both Antarctic and Arctic chemical ozone depletion is caused by elevated chlorine in the atmosphere $^{8-14}$ , the inescapable conclusion is that the depletion is caused by continued use of the CFCs.  $\square$

Received 31 July; accepted 28 December 1995.

- Prather, M. J. & Watson, R. T. *Nature* **344**, 729–734 (1990).
- Elkins, J. W. et al. *Nature* **364**, 780–783 (1993).
- Gunson, M. R. et al. *Geophys. Res. Lett.* **21** (20), 2223–2226 (1994).
- Zander, R. et al. *J. geophys. Res.* **95** (D12), 20519–20526 (1990).
- Zander, R., Gunson, M. R., Foster, J. C., Rinsland, C. P. & Namkung, J. J. *atmos. Chem.* **15**, 171–186 (1992).
- Mankin, W. G. & Coffey, M. T. *J. geophys. Res.* **88**, 10776–10784 (1983).
- Rinsland, C. P. et al. *J. geophys. Res.* **96**, 15523–15549 (1991).
- Solomon, S. *Nature* **347**, 347–354 (1990).
- Anderson, J. G., Toohey, D. W. & Brune, W. H. *Science* **251**, 39–46 (1991).
- Brune, W. et al. *Science* **252**, 1260–1266 (1991).
- Waters, J. et al. *Nature* **362**, 597–602 (1993).
- Manney, G. L. *Nature* **370**, 429–434 (1994).
- Bojkov, R. D., Zerefos, C. S., Balis, D. S., Ziomass, I. C. & Bais, A. F. *Geophys. Res. Lett.* **20** (13), 1351–1354 (1993).
- Muller, R. et al. *J. geophys. Res.* (in the press).
- World Meteorological Organization Rep. 20, vol. 1 (1990).
- Lober, J. M. et al. in *Global Biomass Burning* (ed. Levine, J. S.) 289–304 (MIT Press, 1991).
- Symonds, R. B. et al. *Nature* **334**, 415–418 (1988).
- Mankin, W. G. et al. *Geophys. Res. Lett.* **19** (2), 179–182 (1992).
- Molina, M. J. & Rowland, F. S. *Nature* **249**, 810–814 (1974).
- Cicerone, R. J., Stolarski, R. S. & Walters, S. *Science* **185**, 1165–1167 (1974).
- Roberts, P. C. *Business Week* **29** (June), 26 (1995).

- Begley, S. *Newsweek*, October, p. 71 (1993).
- Russell, J. M. et al. *J. geophys. Res.* **98** (D6), 10777–10797 (1993).
- Russell, James M. *J. geophys. Res.* (in the press).
- Russell, James M. et al. *J. geophys. Res.* (in the press).
- Cicerone, R. J. *Rev. Geophys.* **19**, 123–139 (1981).
- Scientific Assessment of Ozone Depletion*, vol. 37 (WMO, Geneva, 1994).
- Report on Concentrations, Lifetimes, and Trends of CFCs, Halons, and Related Species* (NASA Ref. Publ. 1339, Washington DC, 1994).
- Montzka, S. A., Myers, S. A., Butler, J. H. & Elkins, J. W. *Geophys. Res. Lett.* **21** (23), 2483–2486 (1994).
- Prinn, R. G. et al. *Science* **269**, 187–192 (1995).
- Butler, J. H., Elkins, J. W., Hall, B. D., Cummings, S. O. & Montzka, S. A. *Nature* **359**, 403–405 (1992).
- Brasseur, G. et al. *J. geophys. Res.* **95**, 5639–5655 (1990).
- Weissenstein, D. K., Ko, M. K. W. & Sze, N. J. *Geophys. Res.* **97**, 2547–2559 (1992).
- Kaye, A. et al. *J. geophys. Res.* **96** (D7), 12685–12881 (1991).
- Zander, R. et al. *J. geophys. Res.* **99** (D8), 16737–16743 (1994).
- Schmidt, U. & Khedim, A. *Geophys. Res. Lett.* **18** (4), 763–766 (1991).
- Bischof, W., Borchers, R., Fabian, P. & Kruger, B. C. *Nature* **316**, 708–710 (1985).

ACKNOWLEDGEMENTS. We thank members of the HALOE Data Processing Team for producing a high-quality data set, G. Brasseur for use of the NCAR 2-D model, and J. Reavis for careful manuscript preparation. R.J.C. and M.L. were supported by NASA.

## Magma mixing by convective entrainment

Don Snyder & Stephen Tait

Institut de Physique du Globe de Paris, Laboratoire de Dynamique des Systèmes Géologiques, 4, place Jussieu, 75252 Paris cedex 05, France

RECENT studies of volcanic eruptions have brought to light a puzzling sequence: the ejection of dense and mixed liquids followed by a larger volume of unmixed, less dense liquid $^{1-4}$ . Although these studies have shown that the mixing is associated with the injection of basaltic liquid into a more silicic magma chamber, a paradox remains. Basaltic magma is denser than silicic magma, and should flow along the floor of the magma chamber. This configuration impedes mixing of the two liquids, and appears incompatible with the early eruption of basaltic liquid from the top of the chamber. Previously proposed mixing processes explain neither the eruption sequence nor the eruption of dense liquid during the time of lowest eruption rate $^{1-8}$ . We report here fluid-mechanical experiments that suggest a solution: a thermal plume forms over the replenishment inlet, dragging basaltic liquid upward by viscous coupling, and producing mixing only in a localized area without affecting the bulk of the chamber.

Although such a mixing sequence is common, we concentrate on the well documented 1991 eruption of Pinatubo $^{1,2}$  as the

compositions of its ejecta are known in the context of a detailed eruption chronology. In this eruption, the compositions of ejecta increased in silica content from andesitic (in the initial dome and tephra emissions) to dacitic (in the paroxysmal and subsequent eruptions). The andesites appear to be roughly a 64:36 heterogeneous mixture of the dacite and blobs and streaks of basaltic magma $^2$ . The characteristic sizes of these blobs and streaks vary from millimetres to centimetres $^1$ , as in other eruptions $^{3,9,10}$ . The dacite appears not to be a mixture. We address three questions. First, how does dense, basaltic liquid get transported to the top of the chamber? Second, why were the first ejecta a mixture, whereas afterwards, during the highest eruption rate, the ejecta were unmixed dacite (a composition less dense than the preceding andesite)? Third, why are mixing scales of the order of millimetres to centimetres?

This combination of observations is compatible with none of the previously proposed mechanisms (although they may take place in other circumstances). Turbulent fountaining during injection $^{11}$  would produce extensive mixing near the base of the chamber and not favour the later eruption of a large volume of unmixed fluid. Draw-up of the basaltic liquid due to the draining of the chamber through a narrow conduit $^5$  would first tap the overlying silicic layer and would entrain the most basaltic liquid during the period of the highest eruption rate. At Pinatubo, mixed magmas were erupted first, whereas only unmixed liquid was erupted during the highest eruption rate $^{1,2}$ . Gravitational overturn of the basaltic liquid by vesiculation $^{1,2,6}$  predicts mixing throughout the chamber and hence does not explain the later, unmixed dacites.

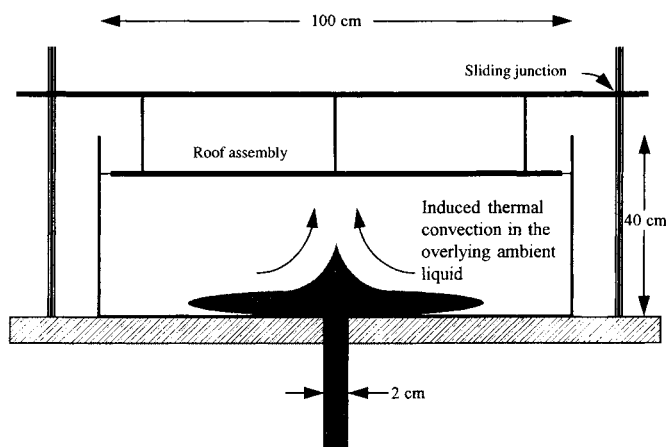


FIG. 1 Schematic diagram of our experimental apparatus. The tank and moveable roof are constructed of acrylic.

In the context of a series of experiments investigating fully developed, multi-layered convection, Huppert *et al.*<sup>7,8</sup> reported a mechanism for magma mixing by viscous coupling, driven by convection rather than magma withdrawal. In this regime, mixing also takes place throughout the overlying layer, and thus does not explain the previously cited eruption sequences. We have, however, done a series of experiments to test whether such a mechanism could be concentrated over the inlet during the transient phase of replenishment and hence affect only a restricted volume of the chamber, thus explaining these puzzling eruption sequences.

We injected hot, dense, low-viscosity liquid upwards into a cooler, less dense, more viscous liquid through a 2-cm-wide by 30-cm-long slot in the tank, as illustrated in Fig. 1. These experiments are intended to provide an analogue of a basaltic magma being fed into a more silicic magma chamber through a dyke. A horizontal roof was permitted to rise with the injection of the dense fluid, mimicking an expanding magma chamber. Both liquids were aqueous solutions of NaCl (to control density) and hydroxyethylcellulose (to control viscosity), so that, like most magmas, the liquids were miscible and had no surface tension between them. In all experiments, the Reynolds number was kept low to avoid turbulence and fountaining.

In previously reported experiments in which the injected and ambient liquids were at the same temperature<sup>12</sup>, the injected liquid spread along the floor of the tank as a viscous gravity current. The upper surface of the current was smooth, and the two

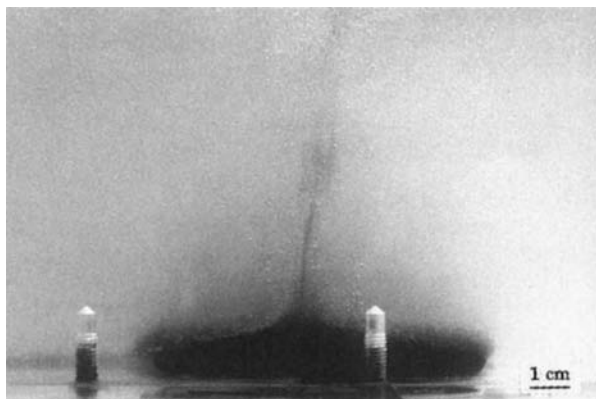


FIG. 2 Photograph of a heated injection experiment showing the development of a thermal plume above the injection slot. The injected liquid is dyed darker than the ambient liquid.

liquids did not mix anywhere along their interface. When the injected liquid was hotter than the ambient liquid, however, the upward flux of heat into the ambient liquid induced thermal convection in the latter. This convection dragged injected liquid upward by viscous coupling, even though the injected liquid was always denser than the overlying liquid (Fig. 2).

Because the heat flux was highest over the inlet (where fresh liquid was constantly entering the system), this was also the site of the most intense thermal plume, and consequently was where most dense injected liquid was dragged upwards. On reaching the roof, the plume spread horizontally to produce a mixed layer there. Secondary plumes often formed toward the distal end of the current, spaced with a wavelength roughly equal to the height of the chamber, but they were always substantially weaker and formed well after the primary plume. In this manner, limited amounts of dense material were locally mixed with the ambient fluid and transported to the top of the tank. After the injection was switched off, the injected liquid was no longer being continually renewed and so the thermal flux over the vent was greatly reduced. Consequently, the entrainment of the dense injected liquid rapidly diminished, eventually ceasing entirely.

This focusing of mixing over the inlet may explain mixing sequences like that observed at Pinatubo. We propose that as basaltic liquid enters a more silicic chamber, a thermal plume over the dyke drags basaltic liquid to the top of the chamber. The initial eruption, perhaps triggered by vesiculation of the more silicic liquid<sup>13</sup> (due to heating), taps this mixed region, producing a heterogeneous andesite. The bulk of the magma, however, remains unmixed, so that the subsequent paroxysmal eruption taps the relatively pristine, dacitic liquid.

Using scaling arguments similar to those used by Solomatov *et al.*<sup>14</sup> in the context of particle entrainment, we derive expressions for the thermal boundary layer thickness ( $d$ ) and the thickness of the sheet of injected liquid that is dragged upward ( $\delta$ ). By balancing the buoyancy forces (the thermal part that drives the convection minus the compositional part that impedes it) and viscous drag, we obtain

$$\alpha \Delta T \rho_{\text{amb}} g d - (\rho_{\text{inj}} - \rho_{\text{amb}}) g \delta \approx \frac{\kappa \eta_{\text{amb}}}{d^2} \quad (1)$$

where  $\alpha$  is the coefficient of thermal expansion of the ambient liquid,  $\Delta T$  is the temperature difference between the injected and ambient liquids,  $\rho_{\text{amb}}$  and  $\rho_{\text{inj}}$  are the densities of the ambient and injected liquids, respectively,  $g$  is the acceleration of gravity,  $\kappa$  is the thermal diffusivity and  $\eta_{\text{amb}}$  is the shear viscosity of the ambient

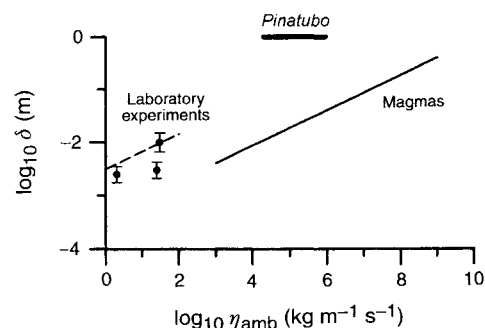


FIG. 3 Plot of the thickness of the entrained sheet of liquid ( $\delta$ ) against the viscosity of the ambient liquid ( $\eta_{\text{amb}}$ ). The predictions of equation (4) are shown for laboratory experiments (broken curve) and for silicate magmas (solid curve). We have held the other independent variables in equation (4) constant because, unlike  $\eta_{\text{amb}}$ , they vary by less than an order of magnitude. Experimental values are shown as solid circles, and these agree well with our order-of-magnitude scaling. The solid horizontal bar at the top of the figure gives the range of viscosity for the ambient silicic magma at Pinatubo, which was estimated from the compositions and phenocryst contents of Pallister *et al.*<sup>2</sup>, the water content of Wallace and Gerlach<sup>15</sup>, and the viscosity models of Shaw<sup>16</sup> (liquid) and Marsh<sup>17</sup> (effect of crystals).

liquid. Solving for  $\delta$ , we find that

$$\delta \approx \frac{\alpha \Delta T \rho_{\text{amb}} d - \frac{\kappa \eta_{\text{amb}}}{d^2 g}}{\rho_{\text{inj}} - \rho_{\text{amb}}} \quad (2)$$

The thermal boundary layer thickness is

$$d \approx \sqrt{\kappa t_i} \approx \left( \frac{\eta_{\text{amb}} Ra_c \kappa}{g \alpha \rho_{\text{amb}} \Delta T} \right)^{1/3} \quad (3)$$

The symbols  $t_i$  and  $Ra_c$  are, respectively, the characteristic initiation time and critical Rayleigh number ( $27\pi^4/4$ ) for the onset of convection. Combining equations (2) and (3) gives

$$\delta \approx 8.7 \left[ \frac{\eta_{\text{amb}} \kappa (\alpha \rho_{\text{amb}} \Delta T)^2}{g (\rho_{\text{inj}} - \rho_{\text{amb}})^3} \right]^{1/3} \quad (4)$$

This scaling corresponds well with our experimental observations (Fig. 3).

Returning to the example of Pinatubo, equation (4) predicts that a thermal plume in the ambient silicic magma ( $10^4 < \eta_{\text{amb}} < 10^6$ )<sup>2,15–17</sup> will drag a sheet of basaltic liquid roughly 1 cm thick to the top of the chamber (Fig. 3). Although we expect that during ascent this sheet will lose heat to the ambient liquid, and possibly mix with it, these effects will not influence our model. This size is consistent with the structures observed in the mixed magmas at Pinatubo. Moreover, the mechanism explains how a small volume of mixed magma was erupted early, followed by unmixed magma.

Although, in general, viscosities of silicic magmas can vary greatly as functions of liquid composition, volatile content and crystallinity, erupted silicic magmas tend not to vary greatly in volatile content; they are usually saturated in volatiles (indeed, volatile saturation most probably triggers the eruptions)—a state that also greatly reduces their viscosity. Consequently, values of  $\eta_{\text{amb}}$  for other volcanic systems are mostly confined between about  $10^3$  and  $10^6$  Pa s. Because of the weak dependence of  $\delta$  on  $\eta_{\text{amb}}$ , equation (4) predicts that sizes of mixing structures will generally be restricted to millimetres to centimetres, as observed<sup>3,9,10</sup>.  $\square$

Received 7 September; accepted 27 December 1995.

1. Pallister, J. S., Hoblitt, R. P. & Reyes, A. G. *Nature* **356**, 426–428 (1992).
2. Pallister, J. S., Hoblitt, R. P., Meeker, G. P., Knight, R. J. & Siems, D. F. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo* (eds Newhall, C. & Punongbayan, R.) (PHIVOLCS, Quezon City, Philippines; Univ. Washington Press, Seattle, in the press).
3. Druitt, T. H., Mellors, R. A., Pyle, D. M. & Sparks, R. S. J. *Geol. Mag.* **126**, 95–126 (1989).
4. Calanchi, N., De Rosa, R., Mazzuoli, R., Rossi, P., Santacroce, R. & Ventura, G. *Bull. Volcanol.* **55**, 504–522 (1993).
5. Blake, S. & Ivey, G. N. *J. Volcan. geotherm. Res.* **27**, 153–178 (1986).
6. Huppert, H. E., Sparks, R. S. J. & Turner, J. S. *Nature* **297**, 554–557 (1984).
7. Huppert, H. E., Sparks, R. S. J. & Turner, J. S. *Earth planet. Sci. Lett.* **65**, 377–381 (1983).
8. Huppert, H. E., Sparks, R. S. J. & Turner, J. S. *J. geophys. Res.* **89**, 6857–6877 (1984).
9. Sigurdsson, H., Carey, S., Palais, J. M. & Devine, J. J. *J. Volcan. geotherm. Res.* **41**, 127–151 (1990).
10. Wolff, J. A. *Geol. Mag.* **122**, 523–640 (1985).
11. Campbell, I. H. & Turner, J. S. *J. Petrol.* **30**, 885–923 (1989).
12. Snyder, D. & Tait, S. *Contr. Miner. Petrol.* (in the press).
13. Sparks, R. S. J., Sigurdsson, H. & Wilson, L. *Nature* **267**, 315–318 (1977).
14. Solomatov, V. S., Olson, P. & Stevenson, D. J. *Earth planet. Sci. Lett.* **120**, 387–393 (1993).
15. Wallace, P. J. & Gerlach, T. M. *Science* **265**, 497–499 (1994).
16. Shaw, H. R. *Am. J. Sci.* **272**, 870–893 (1972).
17. Marsh, B. D. *Contr. Miner. Petrol.* **78**, 85–98 (1981).

ACKNOWLEDGEMENTS. We thank R. Kerr and J. Pallister for reviews.

## Influence of sea-floor spreading on the global hydrothermal vent fauna

Verena Tunnicliffe\* & C. Mary R. Fowler†

\*School of Earth and Ocean Sciences, and Department of Biology, University of Victoria, Victoria V8W 2Y2, Canada

†Geology Department, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK

ONE remarkable discovery of recent decades is the presence of hundreds of unusual species, including fourteen new families, at hydrothermal vents. These animals, unknown from other habitats, live in extreme chemical and thermal conditions around vents on spreading centres of the mid-ocean ridges and back-arc basins. Chemosynthesis provides an *in situ* energy source for the thriving vent fauna. This habitat has existed through the Phanerozoic<sup>1,2</sup> and probably since the Archaean, thus providing sites for long-term adaptation. We now test the hypothesis that animal distribution among hydrothermal vents is related to tectonic plate history<sup>3,4</sup>. The predominant migration pathway is most likely to occur along mid-ocean ridges rather than by shortest oceanic routes. Similarity analyses suggest that the distribution patterns of today's vent fauna display the strong imprint of the timing and geometry of ancient plate boundaries. Study of past ridge geometry provides a method to predict relationships among vent communities yet to be discovered.

Interpretation of modern biogeographic distributions depends upon several factors such as tectonics, catastrophic episodes, ecological requirements and dispersal. Hydrothermal vent biogeography can be examined fairly simply because the communities are strikingly different from other marine assemblages. Present vent communities are distinct from surrounding deep-sea communities and only a few animals inhabit both. Since the first discovery of hydrothermal vent fauna in the Galapagos Rift in 1977, of the 375 recorded species, only 7% occur outside the vent habitat. Fourteen new families constitute 13% of all families

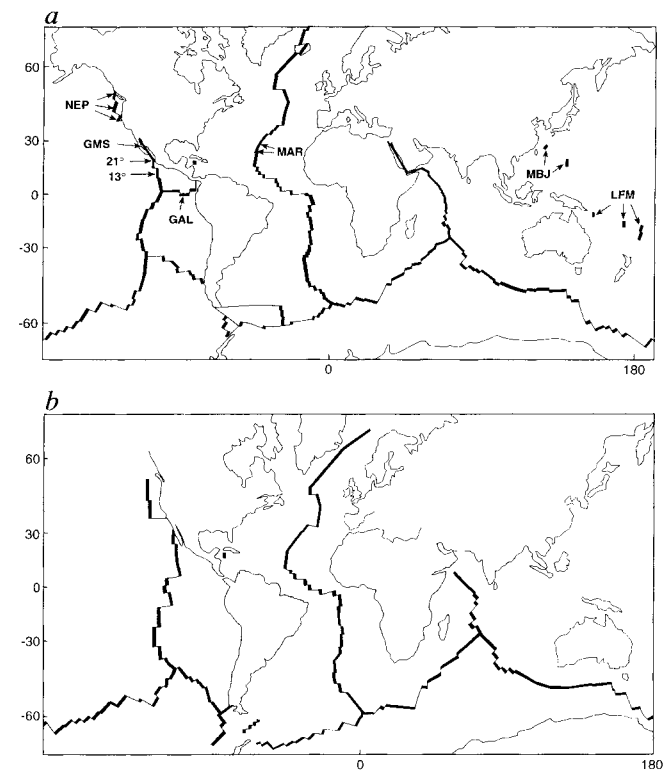


FIG. 1 a, Positions of major biological collections from hydrothermal vents. Known vents are unequally spaced on the globe, reflecting the scientific exploration effort. b, Mid-Cenozoic (23 Myr ago) mid-ocean ridge geometry. Mid-ocean ridges (thick lines) and transform faults (single lines) based on Scotese *et al.*<sup>27</sup>.

recorded at vents. Some taxonomic affinity with sulphur and methane cold seeps on continental margins is evident in the additional seven new families present at both vents and seeps; one recent study on a shared subfamily identifies separate vent and seep clades<sup>5</sup>. As Palaeozoic vent faunas are known<sup>1,2</sup> and some



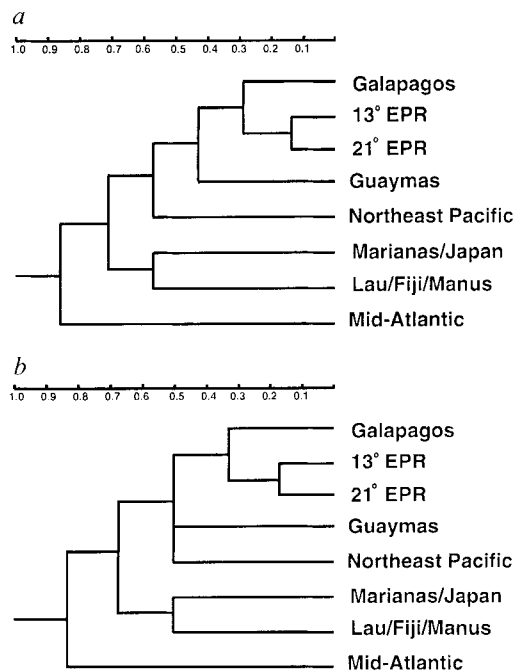


FIG. 2 Cladograms: consensus results of similarity trees generated by three clustering methods (UPGMA, Complete, Single) used in the Jaccard Similarity Coefficient of the NTSYS program. Strict Consensus option was chosen. Several other coefficients were examined in both the NTSYS package as well as a data conversion for phylogenetic equivalent analysis in the PAUP programme. No fundamental pattern differences were detected with other coefficients. Consensus indices ( $C_i$ ) represent the measure of agreement among the trees used to construct the average tree presented here. *a*, Similarity based on species ( $C_i = 1.00$ ). In a separate analysis of sites within the northeastern Pacific, along Juan de Fuca/Explorer/Gorda Ridges, the high similarity reflected regional consistency (not shown). Placement of Guaymas Basin reflects geographical proximity poorly. Habitat type must be important: the ridge is bare at 13° and 21° EPR but heavily sedimented at Guaymas. *b*, Similarity based on genera ( $C_i = 0.83$ ). Use of genera reveals older relations among regions.

taxa have Palaeozoic affinities<sup>6-8</sup>, part of today's vent faunas may be a relict descendant from a Palaeozoic assemblage that inhabited the Pacific/Palaeo-Tethyan mid-ocean ridges.

The taxonomic database from major vent sites (Fig. 1) contains 525 records of published species<sup>9-12</sup> plus a few in description. Unidentified species in described genera were included for analysis at the above-species level. As records are sparse for some sites, those within geographical provinces are combined (Table 1). A presence/absence data set was examined for similarities at species, genus and family level. The number of metazoan phyla is small and many common deep-sea groups are not present, perhaps as a result of physiological barriers<sup>4,8</sup>. Molluscs, polychaetes and arthropods account for 93% of the species. The low species/genus ratio, 1.7 overall, reflects numerous monospecific genera and families. For example, despite great morphological variation and habitat differences, all northeast Pacific tubeworms are the same species<sup>13</sup>. Species proliferation has occurred in a few new genera represented at most vents: the new copepod genus *Stygiopontius*, which has 17 species, is found at every vent site<sup>14</sup>.

The pattern of the species cladogram (Fig. 2a) is unaffected by choice of clustering algorithm. Five amphi-Pacific species are known (Table 2) and all are vent endemic. One, the polychaete *Amphisamytha galapagensis*, features nothing unusual in its reproductive characteristics to suggest great dispersal capability; however, the extensive range of niches it occupies suggests its broad

ecological adaptations may allow a broad distribution<sup>15</sup>. The copepod *Stygiopontius pectinatus* has a noteworthy occurrence at vent fields in two oceans: at the Mid-Atlantic Ridge<sup>14</sup> and Mariana Back-Arc Basin<sup>16</sup>.

Analyses of species similarities ignore information above the species level; older relationships are reflected in higher taxonomic levels. Figure 2b presents a consensus tree of site similarity based on 220 genera. The general structure reflects the species tree well. Nearly all nodes of the cladogram are robust. However, several other coefficients as well as a weighted technique that includes both genera and species, identified a highly variable branch: placement of the sedimented Guaymas Basin (Gulf of California). Guaymas Basin shares several genera with two sedimented sites on the Juan de Fuca Ridge (Middle Valley and Escanaba Trough). Whereas Guaymas Basin is much closer geographically to East Pacific Rise sites, 21° N and 13° N vents are on bare rock. Both species and genus analyses thus demonstrate the importance of habitat in biogeographical patterns. The Atlantic vent fauna has not evolved completely in isolation: 40% of Atlantic genera, of which half are endemic to vents, are also found in the Pacific. Van Dover<sup>17</sup> postulates that the local environment as controlled by ridge spreading rate may be important here.

What process do the cladograms reflect? If larval dispersal is the main factor determining present species ranges, then present-day site proximity should be important. A correlation analysis was

TABLE 1 Numbers of taxa recorded at sampled vent sites

| Province          | Site                                | Species | Genera | Families | Phyla |
|-------------------|-------------------------------------|---------|--------|----------|-------|
| Eastern Pacific   | Galapagos (GAL)                     | 74      | 57     | 38       | 7     |
| Eastern Pacific   | 13° N, EPR (13°)                    | 100     | 79     | 48       | 8     |
| Eastern Pacific   | 21° N, EPR (21°)                    | 88      | 61     | 37       | 5     |
| Eastern Pacific   | Guaymas Basin (GMS)                 | 41      | 35     | 25       | 4     |
| Northeast Pacific | Juan de Fuca, Explorer, Gorda (NEP) | 73      | 64     | 45       | 6     |
| Western Pacific   | Marianas, Okinawa (MBJ)             | 37      | 35     | 25       | 6     |
| Western Pacific   | Lau, Fiji, Manus (LFM)              | 76      | 58     | 41       | 4     |
| Atlantic          | Mid-Atlantic Ridge (MAR)            | 34      | 27     | 24       | 6     |
| All sites         | World                               | 375     | 220    | 107      | 9     |

Western Pacific sites are back-arc basins. Main taxonomic groups not included are nematodes and ciliates. Numbers of species may reflect collection effort, but not entirely. Over twenty-five submersible cruises to Juan de Fuca have sampled vents in thirteen vent fields on eight ridge segments. The same number of species was collected in six cruises at two major fields of Galapagos Rift.

TABLE 2 Shared vent taxa among sites

|     | GAL | 13° | 21° | GMS | NEP | MBJ | LFM | ATL |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| GAL | ●   | 40  | 30  | 15  | 21  | 13  | 11  | 4   |
| 13° | 43  | ●   | 48  | 16  | 20  | 13  | 17  | 7   |
| 21° | 42  | 54  | ●   | 15  | 21  | 13  | 14  | 2   |
| GMS | 13  | 11  | 11  | ●   | 16  | 10  | 8   | 2   |
| NEP | 7   | 3   | 5   | 6   | ●   | 14  | 11  | 5   |
| MBJ | 3   | 4   | 3   | 1   | 1   | ●   | 19  | 7   |
| LFM | 3   | 2   | 2   | 1   | 2   | 9   | ●   | 6   |
| ATL | 0   | 0   | 0   | 0   | 0   | 1   | 0   | ●   |

Below diagonal: numbers of shared species among sites. Above diagonal: numbers of shared genera among sites. Five species are known across the Pacific: two polychaetes, a limpet, a mite and a copepod. The one shared species with the Atlantic is a copepod.

performed between the species-similarity matrix and the distance between sites by applying the Mantel test of matrix comparison<sup>18</sup>. Two distance measures were used: (1) shortest oceanic distances in the modern ocean not using continental shelf passages; and (2) distances via mid-ocean ridges following ridges and faults that could include shortest distance between disjunct ridges<sup>19</sup>. Because of the disparity of inter-ocean versus intra-ocean distances involved, a stringent significance level was chosen (0.01). Shortest distance was not significant ( $P = 0.018$  for all species and  $P = 0.015$  using endemic species only) but mid-ocean ridge distance was significant ( $r = -0.61$ ; Mantel  $t = -2.4$ ;  $P = 0.008$  for all species and  $P = 0.006$  using endemic species). Thus ridge pathways seem to be key to the patterns. Southern Juan de Fuca sites are closer to 21° N East Pacific Rise by shortest distance than is the Galapagos site, but similarity is much higher between the latter two sites. (Similarity GAL/21° is 0.350 versus NEP/21° at 0.032). This suggests that the taxonomic similarities reflect distance along the ridge system and the main discontinuity caused by the San Andreas Fault (Fig. 1).

The cladograms imply ancestral ties among eastern Pacific sites along the mid-Tertiary Pacific-Farallon ridge. This ridge split into the Juan de Fuca and East Pacific Rise about 30 Myr ago when it intersected the subduction zone along the west coast of North America (Fig. 1b)<sup>3,20</sup>. The more recent start of Galapagos Rift spreading about 20 Myr ago may explain the lower diversity there compared to East Pacific Rise. The history of the Pacific region is complex and subduction has removed palaeomagnetic evidence. The present trans-Pacific ridge pathway is south, via the Pacific-Antarctic Ridge (Fig. 1a), but more direct trans-Pacific ridge pathways existed in the early Cenozoic. The Kula-Pacific ridge<sup>50</sup>, the western end of which may have intersected the Japan trench<sup>21</sup>, provided an east-west connection across the north Pacific until about 55 Myr ago. The present Central Basin ridge in the Philippine Sea was probably an early Tertiary ridge/transform plate boundary between the Pacific and (now subducted) North New Guinea plates<sup>22</sup>. The western end of this ridge was then

trapped behind a new subduction zone at 43 Myr as the Philippine plate formed. These early to mid-Cenozoic trans-Pacific ridge connections may explain the high degree of similarity at generic (Table 2) and familial levels between the northeastern Pacific and western Pacific vents despite their present separation. Larval stepping-stones such as hotspots<sup>23</sup> and whale carcasses<sup>24</sup> are proposed to aid distribution of occasional taxa but the history of ridge activity seems to provide the framework for pattern prediction.

Plate tectonic reconstructions are in agreement with the present faunal relationships which do not support a recent mid-ocean ridge connection from Pacific to Atlantic. An open seaway existed at Panama until closure about 3 Myr ago, but there was no ridge connection. Whereas a mid-Cretaceous plate boundary pathway may have existed through the Caribbean, that connection would have included transform faults and a subduction zone as well as a ridge<sup>25</sup>. A Mediterranean pathway from the western Pacific via the Tethyan mid-ocean ridge system was broken by long continental faults and so would not have been possible<sup>26</sup>. Examination of fauna from the southern oceans would test the nature of a recent pathway from the Pacific to the south Atlantic via the Chile Rise and around Cape Horn. The central Mid-Atlantic Ridge formed first in late Jurassic followed by the southern Mid-Atlantic Ridge in early Cretaceous. The northern Mid-Atlantic Ridge formed in late Cretaceous and the Reykjanes Ridge in early Tertiary. Present relationships predict that the (as yet unknown) fauna of the Indian Ocean ridges will reveal the Pacific connection to the Atlantic either north of Australia (mid-late Cretaceous) or south of Australia (late Cretaceous onward).

We suggest that the biogeographical patterns of hydrothermal vent fauna can be explained by the development of the Mesozoic and Tertiary mid-ocean ridges. Testing by study of vent fauna from the Indian, South Atlantic and South Pacific mid-ocean ridges is necessary. Work on phylogenetic versus biogeographical relationships within taxonomic groups of these unusual animals will provide additional tests of distribution mechanisms. □

Received 7 February; accepted 24 November 1995.

1. Banks, D. A. *Nature* **313**, 128–131 (1985).
2. Kuznetsov, A. P., Zaikov, V. V. & Masiennikov, V. V. *Izvestiya Akademii Nauk SSSR Seriya Biologicheskaya* **2**, 232–242 (1990).
3. Tunncliffe, V. *Proc. R. Soc. Lond. B* **233**, 347–366 (1988).
4. Van Dover, C. L. *Trends Ecol. Evol.* **5**, 242–246 (1990).
5. Craddock, C. et al. *Mar. Biol.* **121**, 477–485 (1995).
6. Newman, W. A. *Bull. biol. Soc. Wash.* **6**, 231–242 (1985).
7. McLean, J. H. *Phil. Trans. R. Soc. Lond. B* **319**, 1–32 (1988).
8. Tunncliffe, V. *Palaios* **7**, 338–350 (1992).
9. Barnard, J. L. & Ingram, C. *Smithson. Contr. Zool.* **499**, 1–80 (1990).
10. Tunncliffe, V. *Mar. Biol. Oceanogr. A. Rev.* **29**, 319–407 (1991).
11. Waren, A. & Bouchet, P. *Zool. scripta* **22**, 1–90 (1993).
12. Desbruyeres, D. et al. *Mar. Geol.* **116**, 227–242 (1994).
13. Southward, E. C., Tunncliffe, V. & Black, M. *Can. J. Zool.* **73**, 282–295 (1995).
14. Humes, A. G. *Bull. mar. Sci.* **41**, 645–788 (1987).
15. McHugh, D. & Tunncliffe, V. *Mar. Ecol. Prog. Ser.* **106**, 111–120 (1994).
16. Humes, A. G. *J. nat. Hist.* **24**, 289–304 (1990).

17. Van Dover, C. L. in *Hydrothermal Vents and Processes* (eds L. M. Parson et al.) Geol. Soc. Lond. spec. Publ. **87**, 257–284 (1995).
18. Douglas, M. E. & Endler, J. A. *J. theor. Biol.* **99**, 777–795 (1982).
19. Tunncliffe, V., Fowler, C. M. R. & McArthur, A., in *Segmentation of Mid-Ocean Ridges* Geol. Soc. Lond. spec. publ. (in the press).
20. Atwater, T. in *The Eastern Pacific Ocean and Hawaii*, Vol. N, *The Geology of North America* (eds Winterer, E. L., Hussong, D. M. & Decker, R. W.) 21–71 (Geol. Soc. Am., Boulder, Colorado, 1989).
21. Osozawat, S. *Geology* **22**, 1135–1138 (1994).
22. Jolivet, L., Huchon, P. & Rangin, C. *Tectonophysics* **160**, 23–47 (1989).
23. Kari, D. M. et al. *Nature* **335**, 532–535 (1988).
24. Bennett, B. A. et al. *Mar. Ecol. Prog. Ser.* **108**, 205–223 (1994).
25. Ross, M. I. & Scotese, C. R. *Tectonophysics* **155**, 139–168 (1988).
26. Geale, W. K. *Tectonophysics* **155**, 285–306 (1988).
27. Scotese, C. R., Gahagan, L. M. & Larson, R. L. *Tectonophysics* **155**, 27–48 (1988).

ACKNOWLEDGEMENTS. We thank L. Franklin and D. Desbruyeres for assistance in data compilation, A. McArthur, E. G. Nisbet, E. Southward and A. Southward for comments, and the DOVE BRIDGE Workshop for refinement of ideas. This research was funded by NSERC Canada.

# Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice

Sony Heru Sumarsono\*, Trevor J. Wilson\*, Martin J. Tymms\*, Deon J. Venter†, Catherine M. Corrick\*, Rihanna Kola\*, Mireille H. Lahoud\*, Takis S. Papas§, Arun Seth|| & Ismail Kola\*¶

\* Molecular Genetics and Development Group, Institute of Reproduction and Development, Monash University, Monash Medical Centre, 246 Clayton Road, Clayton, Victoria 3168, Australia

† Peter MacCallum Cancer Institute, St Andrews Place, East Melbourne, Victoria 3002, Australia

§ Center for Molecular and Structural Biology, Hollings Cancer Center, Medical University of South Carolina, Charleston, South Carolina 29425-2213, USA

|| Department of Pathology, Faculty of Medicine and Faculty of Dentistry, University of Toronto and Women's College Hospital, Toronto, Ontario M5S 1A8, Canada M5S 1A8

**EXPRESSION of *Ets2*, a proto-oncogene<sup>1</sup> and transcription factor<sup>2-5</sup>, occurs in a variety of cell types<sup>6</sup>. During murine development it is highly expressed in newly forming cartilage, including in the skull precursor cells and vertebral primordia<sup>7</sup>. *Ets2* is located on human chromosome 21 (ref. 8) and is overexpressed in Down's syndrome (trisomy 21)<sup>9</sup>. Here we generate transgenic mice to investigate the consequences of overexpression of *Ets2*. We find that mice with less than 2-fold *Ets2* overexpression in particular organs develop neurocranial, viscerocranial and cervical skeletal abnormalities. These abnormalities have similarities with the skeletal anomalies found in trisomy-16 mice and humans with Down's syndrome, in which the gene dosage of *Ets2* is increased<sup>10,12</sup>. Our results indicate that *Ets2* has a role in skeletal development and implicate the overexpression of *Ets2* in the genesis of some skeletal abnormalities that occur in Down's syndrome.**

The constructs shown in Fig. 1 were used to generate three founder lines that overexpressed *Ets2* (Fig. 1b–d) and gave progeny with a characteristic phenotype (Fig. 2a). *Ets2* transgenic mice had shortened snouts, abnormally shaped heads, shorter necks, and kyphosis (Fig. 2a). X-ray analysis showed that the calvaria had not completely developed. Furthermore, the basilar cartilaginous bones appeared hypoplastic and the cephalic index was significantly greater ( $P < 0.01$ ).

To investigate these abnormalities further, we stained the skeletons of *Ets2* transgenic mice with alizarin red and alcian blue<sup>12</sup>. The crania of transgenic mice were abnormally shaped (Fig. 2a–d) and the calvaria were thinner and hypoplastic. Hypoplasia of the face also occurred (Fig. 2b, c). Furthermore, the intramembranous cartilaginous deposition around the sutures were absent in the skulls of *Ets2* transgenic mice (Fig. 2c). The atlas and axis vertebrae were abnormal and these mice had shorter necks (Fig. 2d). These findings demonstrate that *Ets2* transgenic mice have abnormalities in skeletal bones derived from intracartilaginous ossification (including the face, base of skull and vertebrae); in sutures between the flat bones believed to be derived from chondroid bone (a cartilaginous intermediate); and in regions of the flat bones formed by a process of direct intramembranous ossification that does not involve cartilage.

Histological analysis further demonstrated abnormalities of the skull and cervical vertebrae. The skull displayed disordered flat-bone formation of both the intramembranous and chondroid-derived components. The intramembranous component showed

discontinuous formation of the bony trabeculae from the connective tissue, uneven mineralization of the trabeculae and irregularity in the thickness of the bone on both external and internal periosteal surfaces (Fig. 2e). Interestingly, the diploë were absent (Fig. 2e). The chondroid bone-derived component showed disorganized suture formation, with irregularities in trabecular thickness, shape and mineralization (Fig. 2f). The cervical vertebrae displayed striking abnormalities (Fig. 2g)—most notably in areas of enchondral ossification. Cartilage plates were markedly thinner, the ossification front irregular, and the proliferating and hypertrophied chondrocyte zones, as well as the zone of provisional ossification, were all thinner than in the controls.

We then compared the skeletal abnormalities of *Ets2* transgenic mice with humans with Down's syndrome and mice with trisomy 16 because: *Ets2* is overexpressed in Down's syndrome<sup>9</sup>; *Ets2* is located on mouse chromosome 16 (refs 11, 13) and trisomy 16 mice show similar characteristics to humans with Down's syndrome<sup>14</sup>; and humans with Down's syndrome also develop neurocranial, viscerocranial and cervical spine abnormalities<sup>15</sup>.

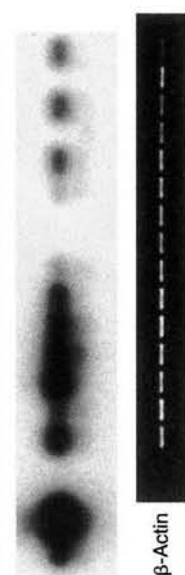
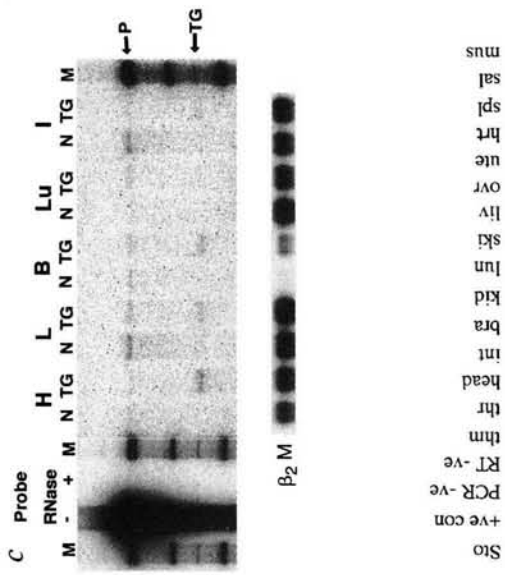
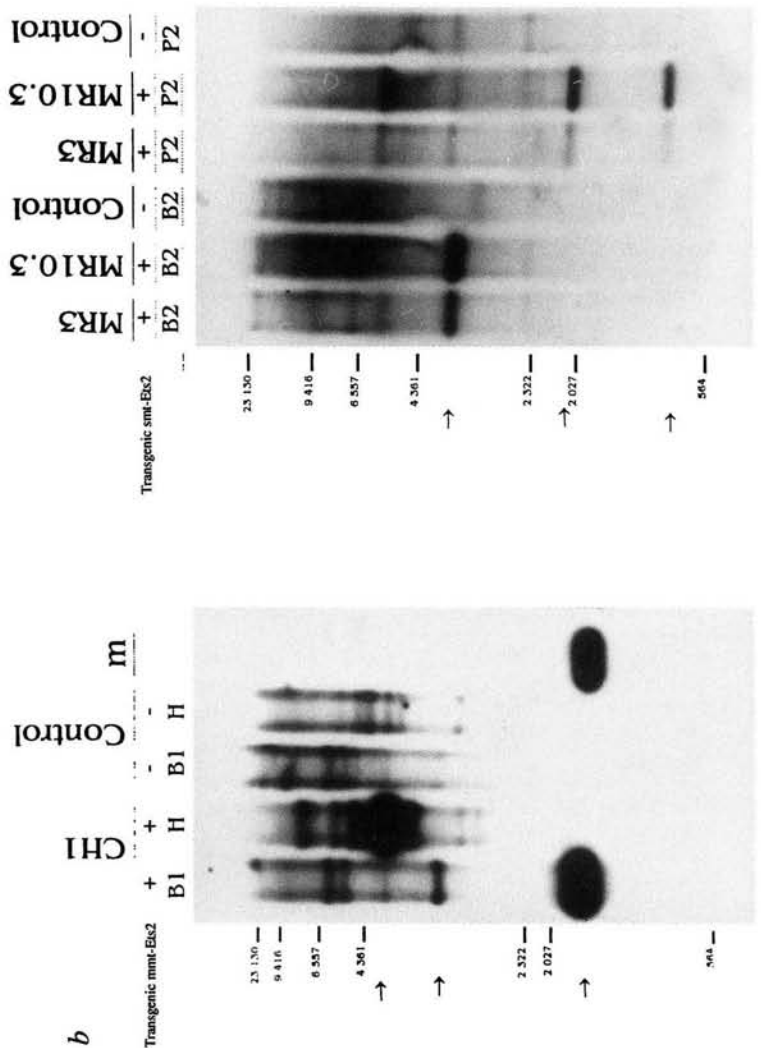
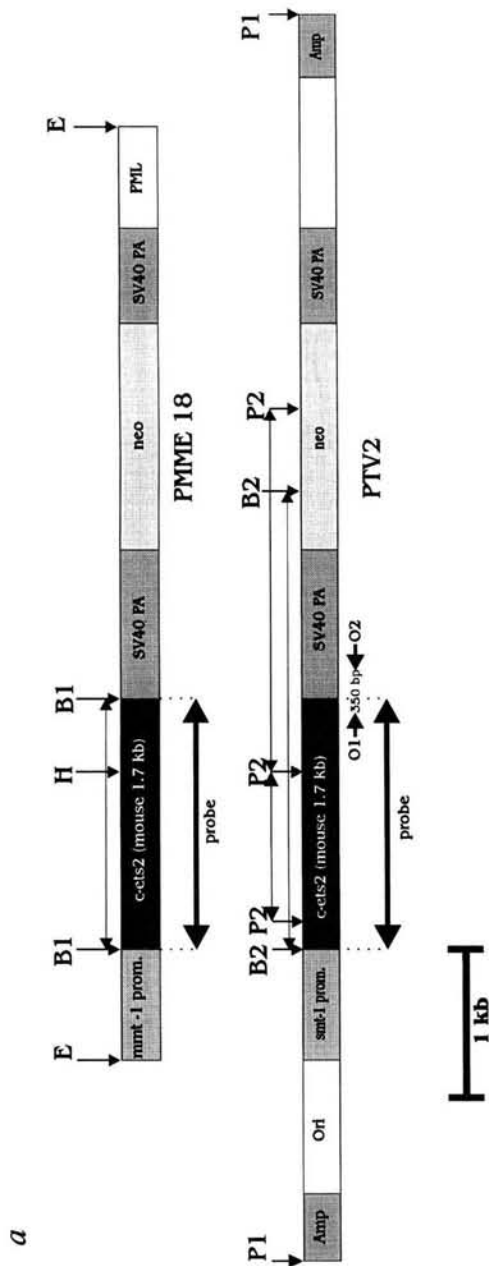
Certain similarities occurred between the *Ets2* transgenic and Down's syndrome phenotypes. Shorter stature and necks occur in Down's syndrome<sup>15</sup>. These individuals have thin calvaria that lack diploic structure and have delayed closure of sutures<sup>15</sup>. Hypoplasia of the cartilage-derived basilar, facial and nasal bones are key features of Down's syndrome. Indeed, it is argued<sup>15</sup> that the characteristic craniofacial features of Down's syndrome are due to abnormal growth and development of the skull. The vertebrae in Down's syndrome show abnormal enchondral ossification with cartilaginous hypoplasia<sup>15</sup>.

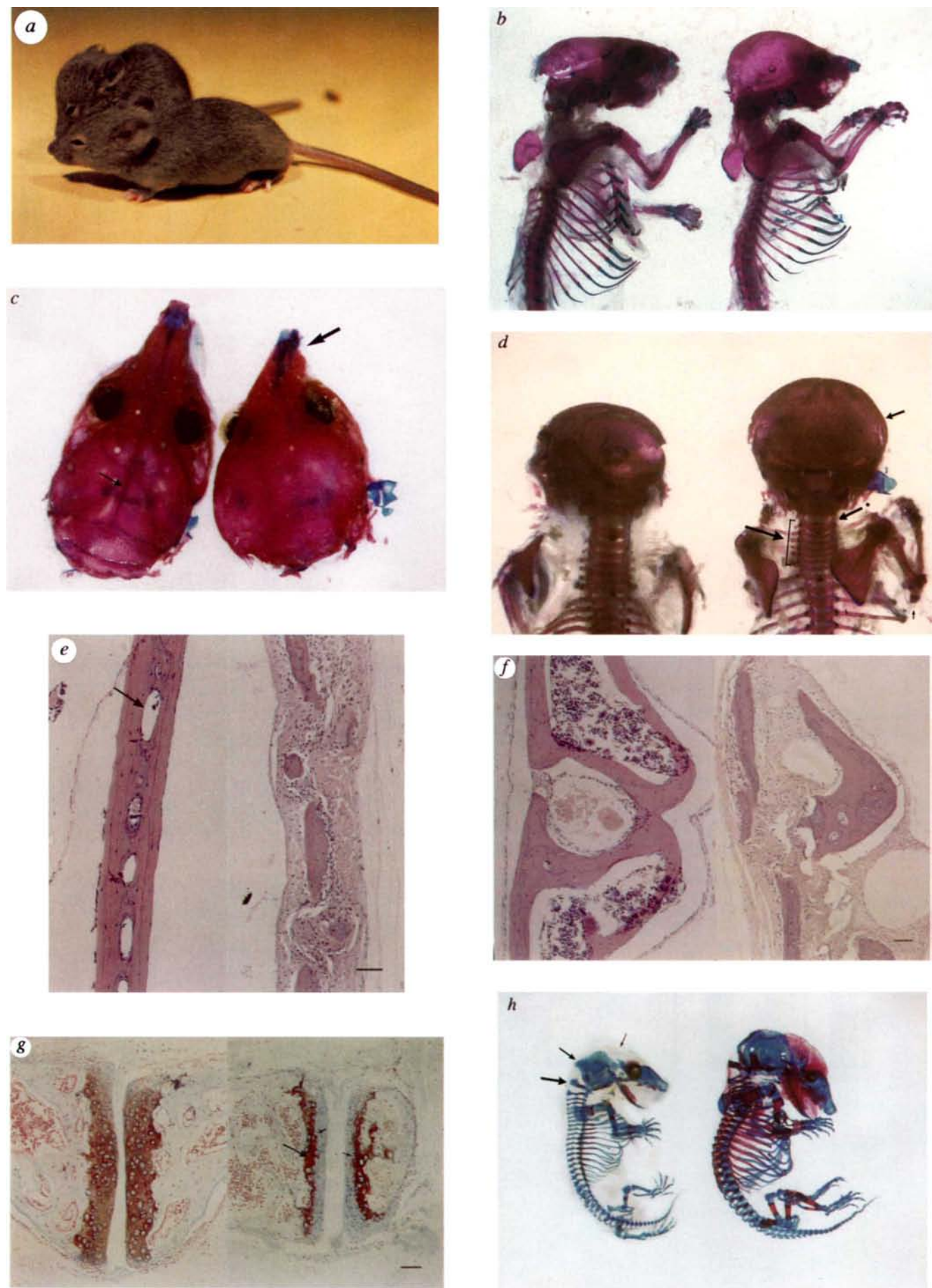
Skeletal anomalies of *Ets2* transgenic mice also resembled those of trisomy-16 mice. Trisomy-16 mice were smaller, had

**FIG. 1** Integration and expression of the transgene. a, Constructs injected<sup>20</sup> were *Ets2* cDNA driven by the mouse metallothionein(*mmt-1*)<sup>1</sup> (top) or the sheep metallothionein(*smt-1*) (bottom) promoter<sup>21</sup>. E, EcoRI; B1, *Bam*HI; H, *Hind*III; B2, *Bgl*II; P2, *Pvu*II; P1, *Pst*I; SV40PA, SV40 poly(A) tail. RT-PCR primer sequences: 01, 5'-GATGGGAGTCAACCTTGC-3'; 02, 5'-CTGTAGGTGTTTGTCCAATATG-3'; b, Southern blots of DNA from transgenic and non-transgenic mice. CH1, *mmt-1-Ets2*; MR3 and MR10.3, *smt-1-Ets2*. Left, *Bam*HI- or *Hind*III-digested DNA hybridized with murine *Ets2* cDNA. Lane m contains marker, *Bam*HI-digested plasmid. Right, *Bgl*II- or *Pvu*II-digested DNA hybridized with *Ets2* cDNA. Arrows, transgene bands. Size calibration (base pairs) is shown on the left. c, Expression of *Ets2* transgene. Top, RNase protection<sup>22</sup> for exogenous *Ets2* and  $\beta_2$ -microglobulin ( $\beta_2m$ ) (control) from 16-day embryonic RNA: head (H; skull with brain removed), liver (L), brain (B), lung (Lu) and intestine (I). Exogenous *Ets2* probe was generated by PCR using primers 01 and 02. The  $\beta_2m$  probe is described in ref. 21. P, undigested full-length probe; TG, the transgene-specific fragment; and M, M<sub>1</sub> markers; N, RNA from non-transgenics; TG, RNA from transgenics. Total *Ets2* mRNA levels (relative to  $\beta_2m$ ) were 1.1–1.8-fold greater in organs of mice constitutively expressing the transgene. Bottom, constitutive expression of transgene by RT-PCR<sup>21</sup> in adult MR3-line mouse with phenotype. Primers 01 and 02 amplify a 350-bp transgene-specific fragment. Same RT-PCR samples were amplified for  $\beta$ -actin (control).  $\beta$ -actin primers: 5'-TCGTGCAACAGGCTCCGG-3' and 5'-GTGGAAGATGATC-GACGCACACCG-3'. Negative controls: PCR-ve (all solutions with no RNA) and RT-ve (RNA with reverse transcriptase). Non-transgenic tissues were consistently negative. Sto, stomach; thm, thymus; thr, thyroid; head, skull with brain removed; int, intestine; bra, brain; kid, kidney; lun, lung; ski, skin; liv, liver; ovr, ovary; ute, uterus; hrt, heart; spl, spleen; sal, salivary gland; mus, muscle. d, Presence of phenotype correlated with transgene expression. In the absence of promoter induction, the percentage of mice with phenotype was: MR3, 24%,  $n = 39$ ; MR 10.3, 30%,  $n = 10$ ; CH1, 11%,  $n = 74$ . Transgene is undetectable or only faintly expressed in organs of mice without phenotype and is induced by ZnSO<sub>4</sub>. Uninduced: brain, testis, spleen and lung from untreated adult mouse with no phenotype; induced: mouse with no phenotype given 5 mg/kg body mass ZnSO<sub>4</sub> at 24 and 12 h before killing. Bottom panel shows  $\beta$ -actin from the same RT-PCR samples.

¶ To whom correspondence should be addressed.







◀ FIG. 2 Phenotype of 4-week-old *Ets2* transgenic mice and 19-day trisomy-16 mouse fetuses. *a*, Transgenic mouse (rear) and control littermate. Note smaller stature, abnormally shaped head and kyphosis. *b*, Alizarin red/alcan blue staining<sup>12</sup> of skeleton of transgenic (right) and control mice (lateral view). Alcan blue stains cartilage and alizarin red, mineralized bone. Note abnormally shaped head and facial hypoplasia. *c*, Staining of skull (top view). Note abnormalities of skull and calvaria; absence of intramembranous cartilaginous deposition (arrowed in control mice—left) and other sutures; hypoplasia of face and nares (larger arrow) and decreased length and increased diameter of skulls. *d*, Staining of skull and cervical skeleton of transgenic (right) and control mice (lateral view). Note abnormally shaped skulls (small arrow); abnormality in atlas-axis vertebrae (larger arrow and asterisk); and shorter neck (large arrow and bracket). *e*, Histology of calvaria of transgenic (right) and controls. Note calvaria are disorganized, irregular in thickness with thinner bone surrounded by increased amounts of connective tissue, and the diploë (arrowed) are lacking. Scale bar, 100 µm. *f*, Histology of the sutures of transgenic (right) and control mice. Sutures are asymmetrical and disorganized; bony trabeculae are irregular in shape, size and mineralization. *g*, Histology of the cervical vertebrae of transgenic (right) and controls. Note markedly thinner cartilage plate (arrows) and irregular ossification front. *h*, Staining of the skeleton of a 16-day trisomy-16 mouse fetus (left) and normal littermate. Note hypoplasia of cartilage that gives rise to parietal and interparietal bones (medium-sized arrows); abnormalities in axis and atlas vertebrae (larger arrow); and delay in bone ossification (small arrow). METHODS. Trisomy-16 mice were generated by mating [*Rb*(6;16); *Rb*(16;17)] mice to wild-type mice; 19-day fetuses were used as these mice are not born alive. Sections were stained with either haematoxylin-eosin or alizarin red-S/alcan blue<sup>23</sup>.

abnormally shaped heads, shorter necks (Fig. 2*h*), cartilaginous hypoplasia, and an alteration in enchondral and intramembranous ossification (Fig. 2*h*). The size of the atlas vertebrae was also smaller (Fig. 2*h*). Furthermore, abnormalities of the intramembranous cartilaginous deposition of parietal skull bones in trisomy-16 mice have been reported<sup>16</sup>. Thus, *Ets2* transgenic mice develop skeletal anomalies reminiscent of those seen in Down's syndrome and trisomy-16 mice. Interestingly, *Ets2* transgenics also had thymus abnormalities, histologically reminiscent of those seen in Down's syndrome (data not shown).

These data have two important implications. First, *Ets2* appears to be important in cartilage/bone development. This is consistent with previous data<sup>7</sup> showing high levels of *Ets2* expression in these tissues during development. Second, overexpression of *Ets2* is implicated in the genesis of skeletal abnormalities that occur in Down's syndrome; this is especially important as one of the major challenges of Down's syndrome research is to identify the relationship between individual genes on human chromosome 21 that induce specific pathologies as part of the Down's syndrome phenotype<sup>17–19</sup>. □

Received 28 September; accepted 30 November 1995.

- Seth, A., Watson, D. K., Blair, D. G. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7833–7837 (1989).
- Wasyluk, B. et al. *Nature* **346**, 191–193 (1990).
- Asciione, R. et al. *Int. J. Oncol.* **1**, 631–637 (1992).
- Seth, A., Robinson, L., Thompson, D. M., Watson, D. K. & Papas, T. S. *Oncogene* **8**, 1783–1790 (1993).
- Thomas, R. S., Tymms, M. J., Seth, A., Shannon, M. F. & Kola, I. *Oncogene* **11**, 2135–2143 (1995).
- Kola, I. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7588–7592 (1993).
- Maroulakou, I. G., Papas, T. S. & Green, J. E. *Oncogene* **9**, 1551–1165 (1994).
- Chumakov, I. et al. *Nature* **359**, 380–387 (1992).
- Kola, I. et al. in *Advances in Human Reproduction* (eds Moeloe, F. A., Affandi, B. & Trounson, A. O.) 309–320 (Casterton-Parthenon, New York, 1995).
- Korenberg, J. R. in *The Morphogenesis of Down Syndrome* (ed. Epstein, C. J.) 43–52 (Wiley-Liss, New York, 1991).
- Reeves, R. H. et al. *Molec. Brain Res.* **2**, 215–221 (1987).
- Kaufman, M. H. in *The Atlas of Mouse Development* 495–507 (Academic, London, 1990).
- Davison, M. T. et al. *Cytogenet. Cell Genet.* **55**, 434–456 (1990).
- Epstein, L. B., Cox, D. R. & Epstein, C. J. *Ann. NY Acad. Sci.* **350**, 171–173 (1980).
- Caffey, J. in *Paediatric X-Ray Diagnosis* Vol. 1, 7th edn 155–157 (Year Book Medical, Chicago, 1978).
- Graus, H., Richtsmeier, J. T. & Oster-Granite, M. L. in *Morphogenesis of Down Syndrome* (ed. Epstein, C. J.) 169–188 (Wiley-Liss, New York, 1991).
- Patterson, D. H. *Sci. Am.* **257**, 42–49 (1987).
- Avraham, K. B., Schickler, M., Sapoznikov, D., Yarom, R. & Groner, Y. *Cell* **54**, 823–829 (1989).
- Reeves, R. H. et al. *Nature Genet.* **11**, 177–184 (1985).
- Iannello, R. C., Young, J., Sumarsono, S., Tymms, M. J. & Kola, I. *Reprod. Fertil. Dev.* **6**, 599–604 (1994).

- Hart, A., Corrick, C. M., Tymms, M. T., Hertzog, P. J. & Kola, I. *Oncogene* **10**, 1423–1430 (1995).
- Tymms, M. J. in *Methods in Molecular Biology—In Vitro Transcription and Translation Protocols* Vol. 37 (ed. Tymms, M. J.) 31–46 (Humana, Totowa, New Jersey, 1995).
- Bancroft, J. D. & Cook, H. C. *Manual of Histological Techniques* (Churchill Livingstone, London, 1984).

ACKNOWLEDGEMENTS. We thank M. Silvestro for technical help, P. Hertzog for discussion, R. Slocombe for preliminary histology preparation, and J. Mison for typing the manuscript. This work was funded in part by grants from The National Health and Medical Research Council (Australia), the Anti-Cancer Council of Victoria and The Medical Research Council (Canada).

## Diet and snake venom evolution

Jennifer C. Daltry, Wolfgang Wüster\* & Roger S. Thorpe

School of Biological Sciences, University of Wales, Brambell Building, Bangor, Gwynedd LL57 2UW, UK

VENOM composition within snake species can show considerable geographical variation<sup>1</sup>, an important consideration because bites by conspecific populations may differ in symptomatology and require different treatments<sup>2–5</sup>. The underlying causes of this phenomenon have never been explained. Here we present evidence that the variation in the venom of the pitviper *Calloselasma rhodostoma* (Serpentes: Viperidae) is closely associated with its diet. We also evaluated other possible causes of geographic variation in venom using partial Mantel tests<sup>6–10</sup> and independent contrasts<sup>11</sup>, but rejected both contemporary gene flow (estimated from geographical proximity) and the phylogenetic relationships (assessed by analysis of mitochondrial DNA) among populations as important influences upon venom evolution. As the primary function of viperid venom is to immobilize and digest prey<sup>12–14</sup> and prey animals vary in their susceptibility to venom<sup>15,16</sup>, we suggest that geographical variation in venom composition reflects natural selection for feeding on local prey.

The Malayan pitviper *C. rhodostoma* is the leading cause of venomous snakebite across much of southeast Asia<sup>3</sup>. Its venom contains many potent digestive enzymes<sup>17</sup> which can cause human victims to suffer severe tissue damage, often leading to permanent deformities among survivors. The precise symptomatology shows some geographical variation<sup>3</sup>, suggestive of variation in venom composition.

For our investigation of the intraspecific variation in venom composition, we collected venom from 67 wild adult *C. rhodostoma* from 36 localities in Vietnam, Thailand, Malaysia and Java (Fig. 1*a*). Each sample was isoelectrically focused (IEF) across polyacrylamide gels containing carrier ampholytes. Ordination analysis of the electrophoretograms revealed strong geographical variation in the venom composition of this species (Fig. 1*b*).

Three hypotheses to account for the geographical variation in *C. rhodostoma* venom were considered. Firstly, variation in venom could be a function of the geographical distance between groups. The opportunity for exchange of venom-coding genes is expected to be higher between spatially close populations and might cause them to produce more similar venom than remote conspecifics. Furthermore, neighbouring populations are more likely to share a similar biotic and abiotic environment, potentially resulting in similar unspecified selection pressures. Secondly, variation in venom may be associated with the patristic phylogenetic relationships among groups. This hypothesis predicts that populations of recent common ancestry produce more similar venoms than populations separated by greater patristic distances<sup>10</sup>. The intra-specific phylogeny of *C. rhodostoma* was reconstructed, as free as possible from perturbation by ecogenetic selection, using restriction-fragment length polymorphism (RFLP) analysis of a

\* To whom correspondence should be addressed.



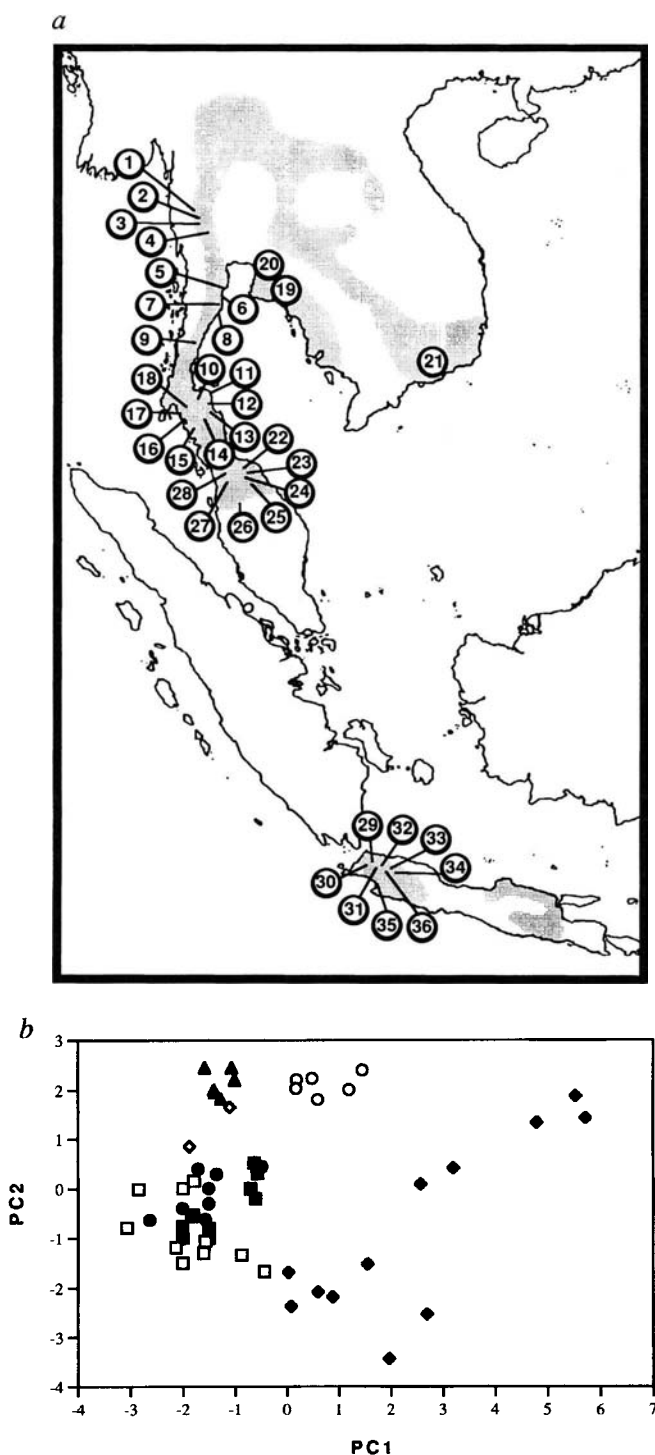


FIG. 1 a, The distribution of 36 sites in Southeast Asia from which venom samples of *Calloselasma rhodostoma* were collected. 67 adult specimens of >400 mm snout-vent length were captured in the Thai provinces of Kanchanaburi (sites 1–4); Prachuap Khiri Khan (5–8); Chumphon (9); Nakhon Si Thammarat (10–14); Krabi (15–18); Rayong (19, 20); also Vung Tau province, Vietnam (21); Kedah state, Malaysia (22–28); West Java, Indonesia (29–36). Shading represents the known distribution range of *C. rhodostoma*. b, Principal components analysis of isoelectrically focused electrophoretograms of venom samples collected in Kanchanaburi (○), Prachuap Khiri Khan (▲), Chumphon (△), Nakhon Si Thammarat and Krabi (●), Rayong (◇), Vietnam (◆), Malaysia (■) and Java (□). Venom was extracted and desiccated within 12 h of capture. The composition of individual samples was compared using isoelectric focusing across precast Ampholine PAGplate gels (Pharmacia), pH range 3.5–9.5. 15 µl of each rehydrated venom (10 mg soluble protein per ml) was loaded and run for 80 min at 1,500 V before staining with Coomassie blue. The influence of gender was removed by excluding bands produced by females only (pI 5.90

767-base-pair (bp) fragment of mitochondrial DNA amplified using polymerase chain reaction (PCR). Thirdly, variation in venom might be associated with geographical variation in diet. Faeces were palpated from each snake upon capture and prey items were identified to taxonomic class. The dietary analysis was supplemented by analysis of the stomach contents and faeces of preserved specimens from the same localities. Adult *C. rhodostoma* eat amphibians, reptiles and endotherms, but the relative contribution of each prey type shows highly significant geographical variation (Fig. 2).

Venom samples were grouped according to geographical origin (Fig. 1a), and thirteen IEF bands varied in distribution among the

TABLE 1 Intercorrelations between venom composition and matrices representing potential causal hypotheses

|                                          | Venom   | Diet    | Geography |
|------------------------------------------|---------|---------|-----------|
| Diet                                     | 0.5957* |         |           |
| Geography                                | 0.4021* | 0.5535* |           |
| Fitch–Margoliash with molecular clock    | 0.2119* | 0.1563  | 0.1342    |
| Fitch–Margoliash without molecular clock | 0.0670  | −0.0354 | −0.0649   |
| Maximum likelihood                       | 0.0649  | −0.0796 | 0.1332    |
| Neighbour joining                        | 0.2346* | 0.3091* | 0.1091    |
| Dollo parsimony                          | 0.0479  | −0.0902 | 0.0946    |

Intercorrelations between distance matrices representing differences in overall venom composition, diet composition, geographical distance, and patristic distances based on phylogenetic trees generated by five algorithms. With the 67 venom samples grouped by geographical origin (Fig. 1a), thirteen distinct bands corresponding to different isoelectric points (pI); Table 2b) varied between the 36 groups. Overall mean venom composition of each group was recorded as a series of numbers denoting the occurrence of each band (so for example, the number 1 indicates that it is present in all group members; 0.5, that it is present in 50%; and 0, that it is absent from all). A 36 × 36 distance matrix representing the variation in overall mean venom composition was compared to similarity matrices derived from three causal hypotheses (see text). Diet was recorded as the percentage of amphibians, reptiles and endotherms (birds and mammals) constituting the diet in each region (Fig. 2). Geographical proximity was computed from the mean latitude and the mean longitude of each group. The phylogenetic relationships among the snakes was assessed using PCR–RFLP analysis of mtDNA. A few millimetres of caudal tissue were removed from all 67 specimens under anaesthesia and stored in 70% ethanol. A 767-base-pair (bp) fragment of mtDNA, chiefly comprising the cytochrome *b* gene (709 bp), was amplified from these biopsies using the PCR oligonucleotide primers (5′) L14841 (ref. 25) and (3′) MVZ16 (ref. 26). Seven restriction endonucleases (*DdeI*, *EcoRI*, *HinfI*, *NciI*, *NlaIII*, *ScrFI*, *TaqI*) identified 15 different haplotypes among these fragments. Phylogenetic trees were generated based on the modal haplotype of each of the 36 regional groups (maximum likelihood and Dollo parsimony algorithms<sup>27</sup>) or on the genetic distances between them, obtained using program DA, REAP<sup>28</sup> (neighbour-joining and Fitch–Margoliash with and without a molecular-clock assumption<sup>27</sup>). The distance-based methods have the advantage that they consider within-population variation. Three main lineages are resolved by all algorithms: Western Thailand (Kanchanaburi and Prachuap Khiri Khan provinces); Vietnam with the rest of Thailand; Malaysia with Java (plus one group from Rayong province in Thailand). Patristic distance matrices were computed between each pair of groups for each of the 5 phylogenetic trees.

\*Indicates significance beyond  $P < 0.05$  after sequential Bonferroni correction to all tests<sup>29</sup>.

and 6.90). Among the 67 venom samples, 53 different types were recognized; snakes that produced identical venoms invariably came from the same locality. Each variant was coded to denote the presence (1) or absence (0) of each variable pI band and a principal components (PC) analysis was conducted upon these data. The first two component scores (PC1, PC2) account for 27.97 and 14.67% of the total variation respectively. Mantel tests were run on matrices representing the totality of the between-population variation in venom composition.

36 groups. Partial Mantel tests were used to evaluate the association between the observed patterns of venom variation among groups (expressed as a taxonomic distance matrix) with patterns predicted by the three hypotheses, while simultaneously partialling out the effects of intercorrelation between the hypotheses (Table 1)<sup>6–10</sup>. Regressing out the patristic distances from the largely selectively neutral molecular phylogeny allows one to test for ecogenetic adaptation free of patristic phylogenetic relationships<sup>10</sup>. The partial Mantel test with overall venom composition (combining all variable bands) as the dependent variable, and geographical proximity, patristic distance and diet as independent variables, found diet alone to be significantly partially correlated (Table 2). When the pattern of variation in each band was tested in turn, at least 5 of the 7–9 bands showing significant association with any of the causal hypotheses were significantly associated with average diet at each locality. Only one band was significantly associated with population phylogeny, casting serious doubt upon the hitherto unchallenged reliability of venom as a taxonomic tool<sup>18–20</sup>. These conclusions are supported by independent contrast methods<sup>11</sup> which show diet and venom to be correlated, free of phylogenetic effects (Table 2).

The variable components have not yet been identified, but preliminary data indicates that the variation in IEF electropho-

retograms is congruent with geographical variation in the venoms' enzymatic activities (unpublished data). Pitvipers typically ingest relatively large prey items<sup>21</sup> and their venoms seem to have two main functions, namely prey immobilization and digestion<sup>12–14</sup>. As prey taxa vary markedly in their susceptibility to snake venom<sup>15,16</sup>, it would seem that natural selection has caused different *C. rhodostoma* populations to produce venoms appropriate for subduing and digesting the local diet. This conclusion is further supported at the within-population level by the simultaneous ontogenetic changes in diet and venom composition.

It should be noted that venoms of captive-bred *C. rhodostoma* produced electrophoretic profiles identical to those of wild specimens from the same locality as their parents, in spite of their unnatural diet in captivity. This indicates that the venom/prey association is inherited rather than environmentally induced; studies of other snakes support the conclusion that venom composition is under strict (non-plastic) genetic control<sup>22,23</sup>.

The findings of this study have important implications for snakebite therapy. The compositional variation elucidated using IEF may well reflect variation in the venoms' immunological properties<sup>24</sup>, and it should be feasible to produce a more widely effective antivenom against *C. rhodostoma* by taking the geographical variation into account<sup>1</sup>. Unfortunately, our results also

TABLE 2 Significance of partial regressions between venom composition and causal hypotheses

|                                             | Geographic proximity | Patristic distance | Diet    |
|---------------------------------------------|----------------------|--------------------|---------|
| (a) Overall venom composition               |                      |                    |         |
| Venom versus geographical distance and diet | 0.2525               |                    | 0.0001* |
| Fitch–Margoliash with molecular clock       | 0.5358               | 0.0225             | 0.0001* |
|                                             |                      | 0.0217             | 0.0002* |
| Fitch–Margoliash without molecular clock    | 0.2255               | 0.1329             | 0.0001* |
|                                             |                      | 0.1435             | 0.0001* |
| Maximum likelihood                          | 0.3792               | 0.1338             | 0.0001* |
|                                             |                      | 0.0905             | 0.0001* |
| Neighbour joining                           | 0.2272               | 0.2490             | 0.0001* |
|                                             |                      | 0.3075             | 0.0001* |
| Dollo parsimony                             | 0.3578               | 0.0976             | 0.0001* |
|                                             |                      | 0.0916             | 0.0001* |
| (b) Individual venom bands†                 |                      |                    |         |
| Band 1, p/ 9.40                             | 0.0576               | 0.7012             | 0.0046‡ |
| Band 2, p/ 9.15§                            | 0.0009*              | 0.0018             | 0.0506  |
| Band 3, p/ 8.55                             | 0.0041               | 0.1680             | 0.1023  |
| Band 4, p/ 8.45                             | 0.6542               | 0.2217             | 0.3542  |
| Band 5, p/ 8.25                             | 0.0083               | 0.6280             | 0.0011* |
| Band 6, p/ 7.85                             | 0.6182               | 0.0862             | 0.4000  |
| Band 7, p/ 7.45                             | 0.8052               | 0.0015*            | 0.7296  |
| Band 8, p/ 6.55                             | 0.0321               | 0.6700             | 0.0012* |
| Band 9, p/ 6.10                             | 0.1113               | 0.6429             | 0.7233  |
| Band 10, p/ 6.00                            | 0.0009*              | 0.0160             | 0.0624  |
| Band 11, p/ 5.80                            | 0.0542               | 0.4829             | 0.0012* |
| Band 12, p/ 5.70                            | 0.0247               | 0.4727             | 0.0002* |
| Band 13, p/ 5.45                            | 0.0044               | 0.4944             | 0.0006* |

a, Partial Mantel matrix association tests: null hypothesis probabilities for the partial regression between overall venom variation and causal hypotheses (diet, geographical distance and each of the 5 phylogenetic trees). Other potential ecogenetic factors, such as rainfall, temperature and vegetation, were tested but showed no association with any venom band. Partial Mantel tests<sup>6–10</sup> show that the pattern of geographical variation in overall venom composition is strongly associated with diet, but not with geographical distance, or phylogeny, irrespective of the phylogenetic algorithm used. To control for possible confounding effects from the significant intercorrelation between the causal hypotheses, further Mantel tests were run that excluded either geographical distance or patristic distance from the analyses. In all cases, venom composition retains its significant partial regression with diet, but not with the other causal hypotheses. Each partial Mantel test entailed 10,000 randomizations. The comparative method of independent contrasts under the gradual change model<sup>11</sup> (as well as other models) supports these conclusions as it gives a significant correlation ( $r = 0.669$ ;  $P < 0.01$ ) between diet and venom, where the continuous variables representing venom and diet variation are the largest principal coordinate of the venom and diet matrices respectively. However, this procedure is less appropriate than the partial Mantel test because it cannot control for possible gene flow between populations (or other multidimensional hypotheses), and does not, on its own, consider multiple competing hypotheses free of distribution. b, Null hypothesis probabilities for the partial correlation between the 13 individual venom components and causal hypotheses. At least 5 of the 7–9 bands showing significant association with any causal hypothesis were significantly associated with diet, two with geographical proximity (probably reflecting the greater opportunity for gene flow between spatially close demes) and only one with patristic distance (phylogeny).

\*Significance beyond  $P < 0.05$  after sequential Bonferroni correction to all tests within the table<sup>29</sup>.

†Probability values for individual bands are based on the Fitch–Margoliash tree with molecular clock assumption. The use of the other phylogenetic trees resulted in identical patterns of significant association except where noted below.

‡Significantly associated with diet when phylogeny is represented by the Dollo parsimony tree.

§Not significantly associated with any causal hypothesis if phylogeny is represented by the maximum likelihood tree.

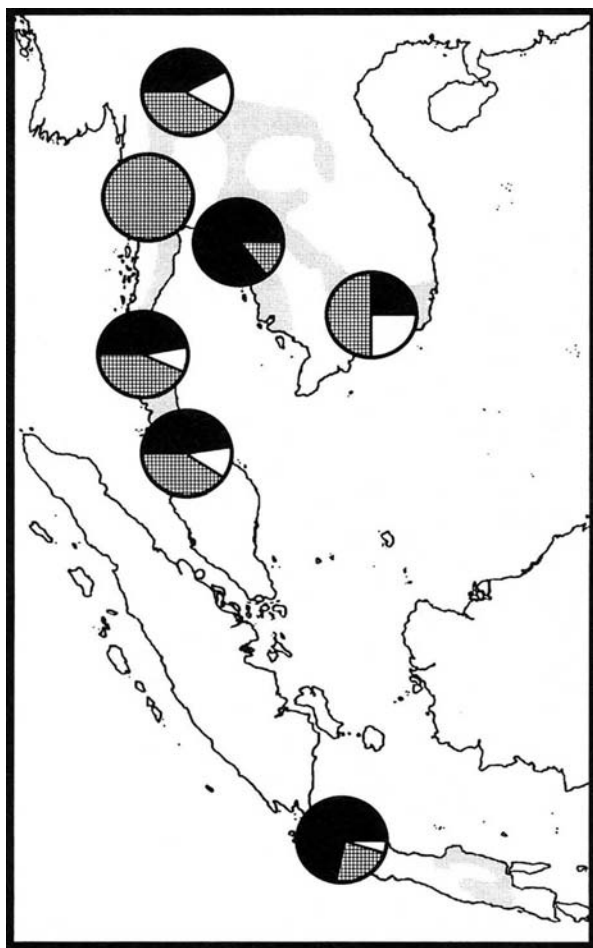


FIG. 2 Geographical variation in the diet of adult *C. rhodostoma*. 216 prey items were identified from faeces and stomach contents. The pie charts represent the contribution of amphibians (white), reptiles (cross-hatched) and endotherms (black) as a percentage of the number of items identified from specimens of >400 mm snout-vent length. Background shading represents the known distribution range of this snake.

indicate that it may not be easy to predict how the venom of other medically important snakes will vary; the population phylogeny may not provide a reliable guide and few species have been subjected to sufficiently rigorous ecological studies to enable patterns of venom variation to be inferred from their diet. □

Received 25 April; accepted 22 November 1995.

1. Chippaux, J.-P., Williams, V. & White, J. *Toxicon* **29**, 1279–1303 (1991).
2. Barrio, A. & Brazil, O. V. *Acta physiol. Lat. Am.* **1**, 291–308 (1951).
3. Warrell, D. A. in *Natural Toxins: Animal, Plant and Microbial*. (ed. Harris, J. B.) 25–45 (Clarendon, Oxford, 1986).
4. Jayanthi, G. P. & Veerabasappa Gowda, T. *Toxicon* **26**, 257–264 (1988).
5. Wüster, W., Otsuka, S., Malhotra, A. & Thorpe, R. S. *Biol. J. Linn. Soc.* **47**, 97–113 (1992).
6. Brown, R. P., Thorpe, R. S. & Baez, M. *Nature* **352**, 60–62 (1991).
7. Thorpe, R. S. & Baez, M. *Biol. J. Linn. Soc.* **48**, 75–87 (1993).
8. Thorpe, R. S., et al. in *Phylogenetics and Ecology* (eds. Eggleston, P. & Vane-Wright, R.) 189–206 (Linnean Soc. Symp. Ser. No. 17, Academic, London 1994).
9. Castellano, S., Malhotra, A. & Thorpe, R. S. *Biol. J. Linn. Soc.* **52**, 365–375 (1994).
10. Thorpe, R. S., Malhotra, A., Black, H., Daltry, J. C. & Wüster, W. *Phil. Trans. R. Soc. Lond.* **B349**, 61–68 (1995).
11. Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).
12. Thomas, R. G. & Pough, F. H. *Toxicon* **17**, 221–228 (1979).
13. Mackessy, S. P. *Copeia* **1988**, 92–101 (1988).
14. Hayes, W. K. *Toxicon* **29**, 867–875 (1991).
15. Minton, S. A. & Minton, S. R. *Venomous Reptiles* (Scribners, New York, 1969).
16. Elliott, W. B. in *Biology of the Reptilia, Physiology B Vol. 8* (eds Gans, C. & Gans, K. A.) 163–435 (Academic, London, 1978).
17. Tan, N.-H. *Tropical Biomedicine* **8**, 91–103 (1991).
18. Tu, A. T. & Adams, B. L. *Nature* **217**, 761–762 (1968).
19. Foote, R. & MacMahon, J. A. *Comp. Biochem. Physiol.* **57B**, 235–241 (1977).
20. Chen, Y., Wu, X. & Zhao, E. *Toxicon* **22**, 53–61 (1984).
21. Pough, F. H. & Groves, J. D. *Am. Zool.* **23**, 443–454 (1983).
22. Boche, J., Chippaux, J. P. & Courtois, B. *Bull. Soc. Path. Exot.* **74**, 356–366 (1981).
23. Gregory-Dwyer, V. M., Egan, N. B., Bianchi Bosio, A., Righetti, P. G. & Russell, F. E. *Toxicon* **24**, 995–1000 (1986).

24. Tu, A. T. & Ganthavorn, S. J. *Herpet.* **12**, 105–107 (1978).
25. Kocher, T. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6196–6200 (1989).
26. Moritz, C., Schneider, C. J. & Wake, D. B. *Syst. Biol.* **41**, 273–291 (1992).
27. Felsenstein, J. *PHYLP* version 3.3 (Univ. Washington, Seattle, 1990).
28. McElroy, D., Moran, P., Bermingham, E. & Kornfield, I. *Restriction Enzyme Analysis Package* version 4.0 (Univ. Maine, Massachusetts, 1992).
29. Rice, W. R. *Evolution* **43**, 223–225 (1989).

ACKNOWLEDGEMENTS. We thank N.-H. Tan and G. Ponnudurai for enzyme assays, J. Norman for advice on molecular methods, B.J.F. Manly and T. Garland for programs, and the museums that loaned specimens to us. This research was primarily funded by a SERC studentship (J.C.D.) and the Leverhulme Trust (R.S.T. and A. Malhotra), with support from NERC (W.W.) and the Royal Society (R.S.T.).

## Dose-response relationships for resetting of human circadian clock by light

Diane B. Boivin\*, Jeanne F. Duffy\*, Richard E. Kronauer\*† & Charles A. Czeisler\*‡

\* Circadian, Neuroendocrine and Sleep Disorders Section, Division of Endocrinology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, Massachusetts 02115, USA

† Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

SINCE the first report in unicells<sup>1</sup>, studies across diverse species have demonstrated that light is a powerful synchronizer which resets, in an intensity-dependent manner, endogenous circadian pacemakers<sup>1–5</sup>. Although it is recognized that bright light (~7,000 to 13,000 lux) is an effective circadian synchronizer in humans<sup>6–10</sup>, it is widely believed that the human circadian pacemaker is insensitive to ordinary indoor illumination (~50–300 lux)<sup>11</sup>. It has been proposed that the relationship between the resetting effect of light and its intensity follows a compressive nonlinear function<sup>12</sup>, such that exposure to lower illuminances still exerts a robust effect<sup>13</sup>. We therefore undertook a series of experiments which support this hypothesis and report here that light of even relatively low intensity (~180 lux) significantly phase-shifts the human circadian pacemaker. Our results clearly demonstrate that humans are much more sensitive to light than initially suspected and support the conclusion that they are not qualitatively different from other mammals in their mechanism of circadian entrainment<sup>14</sup>.

Eight healthy male volunteers (mean ± s.d., 24.50 ± 2.20 years) were exposed to a stimulus consisting of 3 consecutive days of 5-h normal indoor room light (~180 lux) centred 1.5 h after the initial fitted endogenous temperature minimum. Core body temperature was recorded at 1-min intervals throughout the protocol. The initial and final phases were assessed from core body temperature series obtained during 30–50 h constant routine procedures carried out before and after the 3-cycle light stimulus. This procedure consists of a regimen of enforced semirecumbent wakefulness in constant dim light (~10–15 lux) and is used to unmask the phase of the endogenous core body temperature cycle (see ref. 15 for more details). Subjects otherwise remained in dim illuminance levels (~10–15 lux) from the start of their initial constant routine to the end of the experiment (except for darkness during each 8-h sleep episode). Results from these subjects were compared with those of 23 male volunteers (mean ± s.d., 23.16 ± 3.55 years) who were subjected to the same experimental protocol but exposed to either ~0.03 lux (darkness control), ~1,260 lux, or ~9,500 lux during each of the 5-h stimulus episodes<sup>15,16</sup> and whose endogenous circadian temperature amplitude was maintained after exposure to the stimulus<sup>17</sup>. These 31 studies, which represent 341 days of laboratory recording conducted over the past 4 years, were prospectively designed to allow

‡To whom correspondence should be addressed.



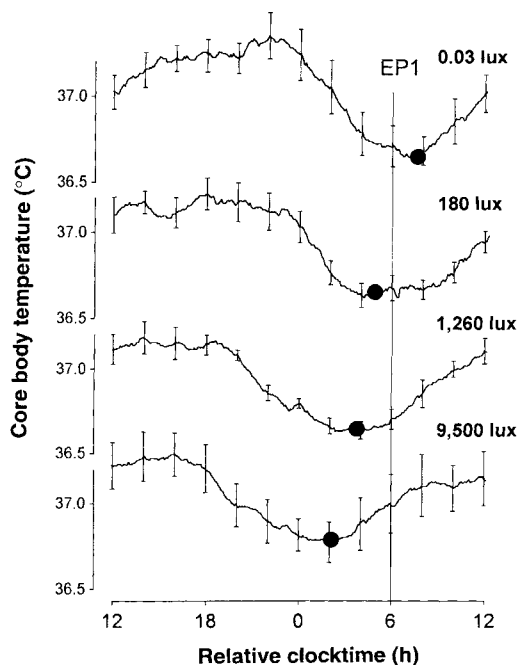
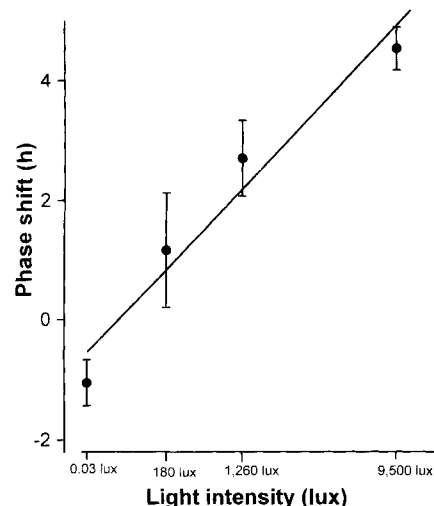


FIG. 1 Average circadian curves of core body temperature observed in each group of subjects after the 3-cycle light or darkness exposure. Results obtained during the second constant routines (CR) were averaged across subjects in each group and plotted (mean  $\pm$  s.e.m.) against relative clocktime. Data collection times were normalized with respect to each subject's initial endogenous circadian temperature minimum (EP1), which was assigned a reference value of 06:00 h and is indicated by the vertical line. The abscissa represents the relative clocktime expressed in h with respect to this initial reference value; the ordinates represent the core body temperature values in degrees Celsius. The filled circles represent the average timing of the final EP observed during the second CR in each group of subjects.

**METHODS.** After 3 baseline days in normal room light ( $\sim 150$  lux during waking hours), subjects remained in dim illumination ( $\sim 10$ – $15$  lux) and the initial phase of the endogenous circadian pacemaker was assessed from core body temperature data obtained during a 30–40 h CR (see ref. 31 for more details), using a dual-harmonic regression model<sup>32</sup>. To exclude masking effects from the previous sleep–darkness episode, the first 5 h of temperature data collected during CR assessments were excluded from analysis. Subjects were then exposed to an initial series of 3-cycle 5-h stimuli at a given intensity, for example,  $\sim 0.03$  lux (darkness),  $\sim 180$  lux,  $\sim 1,260$  lux, or  $\sim 9,500$  lux ( $n = 7$ – $8$  per intensity). The stimulus was centred 1.5 h after the initial fitted endogenous temperature minimum. Light was administered using ceiling-mounted banks of cool-white fluorescent lamps (North American Philips Lighting Corp., Bloomfield, New Jersey, USA). Every 5-h light stimulus was bracketed by 25-min transitional periods<sup>15</sup>.

FIG. 2 Dose–response curve of the phase-shifting effect of 3 cycles of 5-h light exposures on the output of the human circadian pacemaker. The ordinate represents the phase shift observed between the first and second CRs in each group of subjects. The abscissa represents the light intensity (lux) plotted on a cube-root scale. By convention, positive phase shifts represent phase advances and negative phase shifts represent phase delays. Group averages ( $\pm$  s.e.m.) are represented by closed circles and vertical bars, respectively. A linear regression model was applied to the individual phase shifts to evaluate the relationship between the cube root of light intensity and its phase-shifting effects (oblique straight line) (d.f. 1,29,  $r = 0.742$ ,  $P < 0.001$ ).



construction of a dose–response curve to light and replication of our preliminary studies<sup>18</sup>.

Phase advances of the endogenous core body temperature cycle were observed in every light-exposed group (Fig. 1), with a mean phase shift ( $\pm$  s.e.m.) of  $+1.16 \pm 0.96$  h,  $+2.69 \pm 0.69$  h and  $+4.49 \pm 0.36$  h after exposure to  $\sim 180$  lux,  $\sim 1,260$  lux and  $\sim 9,500$  lux, respectively. A delay of  $-1.05 \pm 0.38$  h was observed in the control group of subjects exposed to  $\sim 0.03$  lux. Two-tail ANOVA, with Student–Newman–Keuls correction for multiple comparisons, revealed that the phase advance shifts in every group of light-exposed subjects were significantly different from the phase delay shift observed in the control group of subjects, and that the phase advances observed for the lower ( $\sim 180$  lux) and brighter illuminance ( $\sim 9,500$  lux) groups were significantly different from each other ( $F_{3,27} = 12.36$ ,  $P < 0.0001$ ). One-tail unpaired *t*-test verified the significance of the difference between the  $\sim 180$  lux and the  $\sim 0.03$  lux groups ( $P < 0.003$ ). The relationship between the phase-shifting effect of light and its intensity followed a nonlinear function, and a statistically significant regression was observed between the cube root intensity of the light stimulus and its phase-shifting effect ( $r = 0.742$ ,  $P < 0.0001$ ) (Fig. 2).

This study is the first to characterize the dose–response curve of photic resetting of the human circadian pacemaker. It clearly demonstrates that exposure to light of even normal indoor intensity ( $\sim 180$  lux) can significantly phase-advance the endogenous component of core body temperature cycle, despite conditions that favoured free-running in the control group of subjects. In this latter group, the delay observed is interpreted as a drift of the endogenous temperature cycle, consistent with the slightly

longer than 24-h period of the human circadian pacemaker<sup>16,19,20</sup>. At the stimulus reference phase chosen for these experiments (1.5 h after the initial temperature minimum), the mathematical model of Kronauer<sup>12</sup> predicts a nearly linear relationship between the strength of the drive on the pacemaker and the phase shift achieved by a light stimulus, even though exposure to  $\sim 9,500$  lux elicits Winfree type 0 resetting<sup>8,21</sup> whereas weaker stimuli generate type 1 resetting<sup>22</sup>. This model postulates that the pacemaker drive is proportional to the one-third power of light intensity, implying that the effect of light is preserved despite marked reduction in the physical strength of the stimulus<sup>13</sup>. The present study is consistent with the observation that light as dim as 100 to 500 lux can significantly reduce the nocturnal secretion of melatonin<sup>23–25</sup> and delay its onset time<sup>26</sup>. Our results are also consistent with the recent observation that photoperiod-responsive changes in human circadian rhythms may be suppressed by regular exposure to artificial room light<sup>27</sup>. In addition, these results imply that the significant phase advances reported in control subjects exposed to 300–500 lux<sup>28</sup> can be accounted for by a direct effect of such moderate light intensity on the human circadian pacemaker. The present study supports the earlier suggestion<sup>14</sup>, that an imposed

16-h:8-h light-dark cycle, using ordinary indoor room light, is sufficient to entrain free-running temperature rhythms in human subjects. Furthermore, it reinforces the conclusion, based on the observation of type 0 resetting<sup>8,17</sup> and of light-induced suppression of endogenous circadian amplitude<sup>29</sup>, that the sensitivity of the human circadian pacemaker to the resetting effects of light falls within the range observed among other animal species. Moreover, the present results have important implications for the practical scheduling of phototherapy in the field, since they imply, as already observed<sup>8</sup>, that competitive sources of lighting, of even normal indoor illuminance, can modulate the resetting effect of a very bright light stimulus. Finally, the present study indicates that exposure to relatively low light intensities, such as those produced by artificial lamps, is probably the principal synchronizer of the human circadian pacemaker to the geophysical environment in modern society, where urban dwellers are generally exposed to light below 1,000 lux<sup>30</sup>. □

Received 5 September; accepted 31 October 1995.

- Hastings, J. W. & Sweeney, B. M. *Biol. Bull.* **115**, 440–458 (1958).
- Chandrasekaran, M. K. & Loher, W. J. *exp. Zool.* **172**, 147–152 (1969).
- McGuire, R. A., Rand, W. M. & Wurtman, R. J. *Science* **181**, 956–957 (1973).
- Takahashi, J. S., DeCoursey, P. J., Bauman, L. & Menaker, M. *Nature* **308**, 186–188 (1984).
- Nelson, D. E. & Takahashi, J. S. *J. Physiol. (Lond.)* **439**, 115–145 (1991).
- Van Cauter, E. et al. *Am. J. Physiol.* **266**, E953–E963 (1994).
- Czeisler, C. A. *Circadian Clocks and their Adjustment*. Chichester (Ciba Foundation Symp. 183) 254–302 (Wiley, 1995).

- Czeisler, C. A. et al. *Science* **244**, 1328–1333 (1989).
- Minors, D. S., Waterhouse, J. M. & Wirz-Justice, A. *Neurosci. Lett.* **133**, 36–40 (1991).
- Honma, K. & Honma, S. *Jap. J. Psych. Neurol.* **42**, 167–168 (1988).
- Wever, R. A. *J. Biol. Rhythms* **4**, 161–185 (1989).
- Kronauer, R. E. *Sleep Res.* (abstract) **16**, 621 (1987).
- Klerman, E. B., Dijk, D.-J., Kronauer, R. E. & Czeisler, C. A. *Am. J. Physiol.* **39**, R271–R282 (1996).
- Czeisler, C. A., Richardson, G. S., Zimmerman, J. C., Moore-Ede, M. C. & Weitzman, E. D. *Photochem. Photobiol.* **34**, 239–247 (1981).
- Boivin, D. B., Duffy, J. F., Kronauer, R. E. & Czeisler, C. A. *J. Biol. Rhythms* **9**, 315–331 (1994).
- Kronauer, R. E., Duffy, J. F. & Czeisler, C. A. *Sleep Res.* **22**, 626 (1993).
- Jewett, M. E., Kronauer, R. E. & Czeisler, C. A. *J. Biol. Rhythms* **9**, 295–314 (1994).
- Allan, J. S., Czeisler, C. A., Duffy, J. F. & Kronauer, R. E. *154th Ann. A.A.S. Meeting* **208**, 101 (1988).
- Campbell, S. S., Dawson, D. & Zuley, J. *Sleep* **16**, 638–640 (1993).
- Czeisler, C. A. et al. *Sleep Res.* (abstract) **24**, 505 (1995).
- Winfree, A. T. *J. theor. Biol.* **28**, 327–374 (1970).
- Kronauer, R. E., Jewett, M. E. & Czeisler, C. A. *J. Biol. Rhythms* **8**, 351–360 (1993).
- Gaddy, J. R., Rollag, M. D. & Brainard, G. C. *J. clin. Endocrin. Metab.* **77**, 1398–1401 (1993).
- Bojkowski, C. J. et al. *Horm. Metabol. Res.* **19**, 437–440 (1987).
- McIntyre, I. M., Norman, T. R., Burrows, G. D. & Armstrong, S. M. *J. pineal Res.* **6**, 149–156 (1989).
- Laakso, M., Hättönen, T., Stenberg, D., Allila, A. & Smith, S. J. *Pineal Res.* **15**, 21–26 (1993).
- Wehr, T. A., Giesen, H. A., Moul, D. E., Turner, E. H. & Schwartz, P. J. *Am. J. Physiol.* **269**, 173–178 (1995).
- Honma, K.-I. et al. *Am. J. Physiol.* **268**, R528–R535 (1995).
- Jewett, M. E., Kronauer, R. E. & Czeisler, C. A. *Nature* **350**, 59–62 (1991).
- Cole, R. J. et al. *J. Biol. Rhythms* **10**, 324–334 (1995).
- Czeisler, C. A. et al. *Science* **233**, 667–671 (1986).
- Brown, E. N. & Czeisler, C. A. *J. Biol. Rhythms* **7**, 177–202 (1992).

ACKNOWLEDGEMENTS. Supported by the National Institute of Mental Health, the National Institute on Aging, the National Aeronautics and Space Administration, the National Heart Lung and Blood Institute, and the General Clinical Research Center Program of the National Center for Research Resources; D.B.B. is supported by the 'Fonds de la Recherche en Santé du Québec' and the Medical Research Council of Canada.

## Resetting of the circadian clock by a conditioned stimulus

Shimon Amir & Jane Stewart

Center for Studies in Behavioral Neurobiology, Department of Psychology, Concordia University, 1455 de Maisonneuve Blvd, Montreal, Québec H3G 1M8, Canada

**ENVIRONMENTAL light is the dominant temporal cue for the entrainment of circadian rhythms. In mammals, light entrains circadian rhythms by daily resetting a pacemaker located in the hypothalamic suprachiasmatic nucleus (SCN)<sup>1,2</sup>. Although it is widely held that phase resetting by light involves cellular elements within the SCN that are uniquely responsive to photic cues<sup>1,3</sup>, we now report that non-photic cues that reliably precede the onset of light can, through associative learning, come to activate these elements. In rats, a neutral non-photic stimulus paired with light in pavlovian conditioning trials was capable of eliciting cellular and behavioural effects characteristic of phase-dependent resetting of the pacemaker by light, the expression of the transcription factor Fos in SCN cells, and phase shifts in free-running activity and temperature rhythms<sup>4–7</sup>. Thus an associative learning process, pavlovian conditioning, provides a means whereby environmental cues that predict light onset can come to mimic the effect of light on the SCN pacemaker and thereby bring about entrainment of circadian rhythms.**

In pavlovian conditioning, a neutral stimulus, the conditional stimulus (CS) is paired with an unconditional stimulus (US) known to induce a specific response. After repeated CS–US pairings, the previously neutral stimulus becomes capable of eliciting responses similar to those originally elicited only by the US. To determine whether non-photic stimuli that come to predict light onset through pavlovian conditioning could mimic the effect of light in the SCN, groups of rats were given 10 explicit pairings (groups (CS–US)) of a neutral stimulus consisting of a 20-min exposure to a gentle air disturbance, the CS, and a 15-min light pulse, the US. For each of these pairings, the onset of the air disturbance preceded the onset of the light pulse by 15 min. The stimuli were presented once daily during the third hour of the

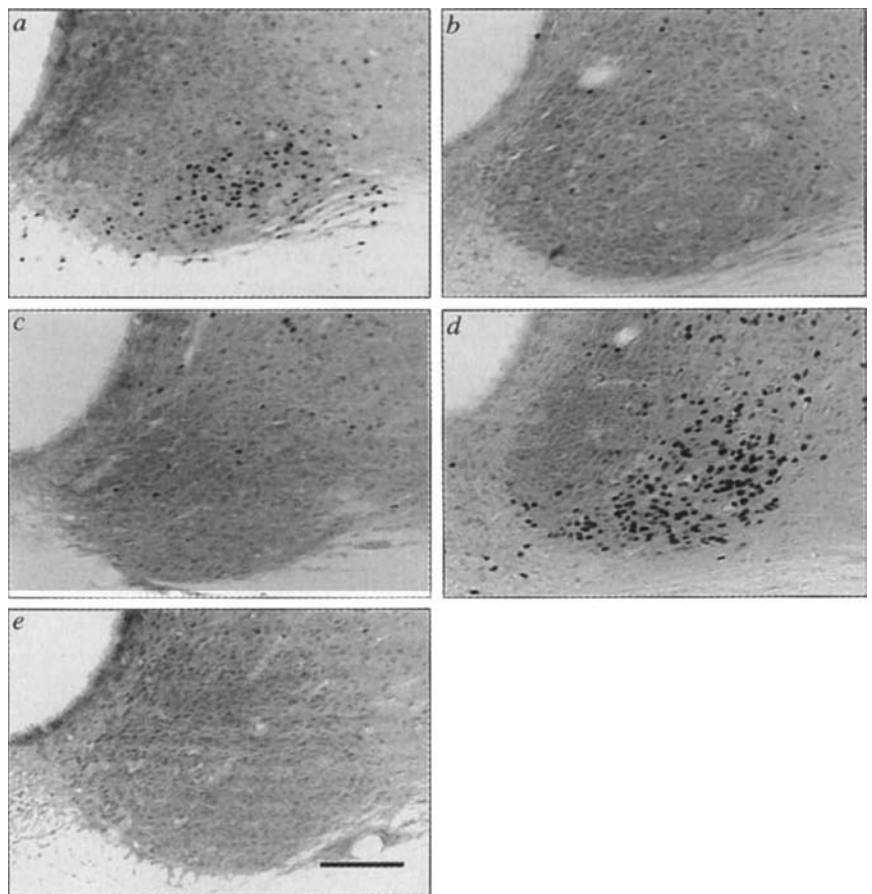
dark phase of the 12h:12h light–dark cycle, a time when light pulses are known to induce both Fos expression in the SCN and phase delays in free-running rhythms<sup>4–7</sup>. Other groups of rats were presented daily with both the CS and the US during the dark phase of the cycle, but the stimuli were explicitly unpaired (groups (CS ≠ US)). On the test for conditioning, all animals were presented with the CS alone. Additional groups were exposed daily to the US alone (groups (US)) as a control for the possible effects of repeated exposure to the US on the subsequent response to the CS (sensitization). Finally, to compare the response to the CS to the response to the US, some animals from group (US) were presented with the US on the test day.

The effect of a 20-min exposure to the CS alone, given 24 h after the last 'conditioning' day, on Fos expression in the SCN is shown in Fig. 1. High levels of Fos immunoreactivity were seen in response to the CS in animals previously given explicit pairings of the air disturbance and the light (group (CS–US)CS in Fig. 1a); in the other groups, presentation of the CS was virtually without effect. Prior exposure to the US alone did not induce a sensitized response to the CS (group (US)CS in Fig. 1c), and particularly noteworthy is the very low level of Fos expression in the SCN in animals in group (CS ≠ US)CS (Fig. 1b). These animals were exposed to both air disturbance and light, similar to that of animals in group (CS–US)CS, but the stimuli were presented non-contingently, indicating that mere experience with these stimuli could not be a factor in the induction of Fos by the CS on test day. Exposure to the US on test day induced strong expression of Fos in the SCN (group (US)US, Fig. 1d). The pattern of distribution of Fos in the SCN induced by the CS in group (CS–US)CS (Fig. 1a) was similar to that induced by the US in group (US)US (Fig. 1d). In these two groups, the largest number of labelled cells was found in the ventrolateral subdivision of the SCN, the principal region of termination of the retinohypothalamic tract<sup>8,9</sup>. No Fos expression was seen in the SCN of animals killed without exposure to either stimulus (Fig. 1e). The mean Fos expression for these groups is shown in Fig. 2.

These data are consistent with the idea that the change in the ability of the CS to induce Fos in the SCN occurred as a result of pavlovian conditioning. Furthermore, they show that a neutral stimulus, through explicit pairing with light, gains access to neural structures and cellular processes in the SCN that are normally activated only by light itself. Next we considered where in the

FIG. 1 Representative photomicrographs of coronal sections through the SCN showing Fos immunoreactivity following exposure to CS or US on the test day in groups. *a*, (CS-US)CS; *b*, (CS  $\neq$  US)CS; *c*, (US)CS; *d*, (US)US. *e*, Fos expression in an untreated control animal maintained in similar conditions and killed in the same period of the dark phase of the cycle. Scale bar, 100  $\mu$ m.

**METHODS.** Male Wistar rats (300–325 g; Charles River, St Constant, QC, Canada) were housed individually in flat black-painted plastic cages (36  $\times$  24  $\times$  20 cm) in temperature-controlled rooms under a 12h:12h light–dark cycle (lights on 20:00–08:00) with free access to food and water. The CS was a 20-min pulse of gentle air disturbance produced by a low-noise fan (Sprite, 1,950 r.p.m.). The US was a 15-min pulse of light produced by a BLB (backlight blue) light source (320–390 nm,  $\lambda_{\text{max}}$  365 nm; Microlites Scientific). The light source provided irradiance of 40  $\mu$ W cm $^{-2}$  at eye level. On the test, animals were presented with either the CS or US and 40 min later were anaesthetized with sodium pentobarbital (100 mg kg $^{-1}$ ) and perfused transcardially with 200 ml cold physiological saline (0.9% NaCl) followed by 400 ml cold, fresh 4% paraformaldehyde in a 0.1 M phosphate buffer (pH 7.3). Control animals were anaesthetized in the dark without prior exposure to either stimulus. Brains were removed, post-fixed in 4% paraformaldehyde overnight (4  $^{\circ}$ C), and 50- $\mu$ m sections were cut by vibratome. Fos immunohistochemistry has been described<sup>7</sup> and used an affinity-purified mouse monoclonal antibody raised against the N-terminal sequence of Fos (corresponding to N-terminal residues 4–17 of human Fos protein; NCI/BCB Repository, Quality Biotech). Fos immunoreactivity was detected with a Vectastain Elite ABC Kit (Dimension Labs) using diaminobenzidine as chromogen.



circadian system both photic and non-photic stimuli might have access to common elements. A prime candidate is the region of the lateral geniculate complex of the thalamus known to be part of the circadian system, the intergeniculate leaflet (IGL)<sup>10–13</sup>. The IGL receives direct input from the retina<sup>14,15</sup> and projects to the SCN via the geniculo–hypothalamic tract<sup>10,11,16–18</sup>. Furthermore, cells in the IGL and in the neighbouring ventral lateral geniculate nucleus (vLGN) respond to non-photic stimuli<sup>19,20</sup>. These observations, combined with the recently described fine synaptic organization within the area, the evidence for axoaxonal contacts, and for the multiple inputs to single intrinsic cells<sup>11</sup>, make the IGL/vLGN well suited for integration of information from photic and non-photic sources. Figure 3 (left panel) shows that exposure to the CS or US on test day induced Fos expression in the IGL/vLGN in all treatment groups. The effect of the CS following conditioning (group (CS-US)CS) was comparable to that of the light itself (see group (US)US). Furthermore, it was significantly greater than that found in groups (CS  $\neq$  US)CS and (US)CS, indicating that explicit CS-US pairings enhanced the response to the CS in this region. Another area of possible integration of the CS and US information is the anterior region of the paraventricular thalamic nucleus (PVT). This midline thalamic structure receives input from the retina and has reciprocal connections with multiple brain areas including the SCN and the IGL<sup>21</sup>. Once again it was found (Fig. 3, right panel) that the response to the CS following conditioning (group (CS-US)CS) was significantly greater than that to the CS in groups (CS  $\neq$  US)CS and (US)CS, and moreover, virtually equalled the response to light itself (group (US)US). Thus, there is evidence for conditioned changes in response to the CS in both the IGL/vLGN and the PVT that parallel the changes in the SCN. Unlike the SCN, these areas show Fos expression in response to both non-photic and photic stimuli before conditioning, suggesting that either could serve as a place for stimulus integration, and, through their projections to the

SCN, provide access for CS information to those elements in the SCN that normally respond only to light.

Having found that this non-photic conditioned stimulus could induce Fos expression in the SCN characteristic of the phase resetting stimulus, light, we next tested the ability of the CS to induce phase shifts in circadian rhythms. Following 'conditioning', animals were placed in constant darkness to allow rhythms to free run; body temperature and activity were monitored continuously.

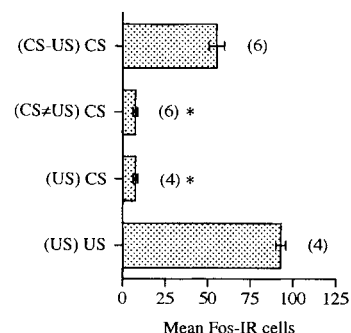


FIG. 2 Mean  $\pm$  s.e.m. number of Fos-immunoreactive (Fos-IR) cells on one side of the SCN in each treatment group (see text) following exposure to the CS or the US on the test day; number of animals per group is shown in parentheses. Serial brain sections through the entire rostral–caudal extent of SCN were scanned using a computerized image analysis system and mean cell count of the five sections on one side of the SCN exhibiting the highest Fos expression were obtained for each animal. Treatment effects were evaluated by a one-way ANOVA ( $F(3, 16) = 156.8$ ,  $P < 0.0001$ ) and individual group differences by Scheffé post hoc tests; asterisk indicates difference from group (CS-US)CS and (US)US,  $P < 0.0001$ .



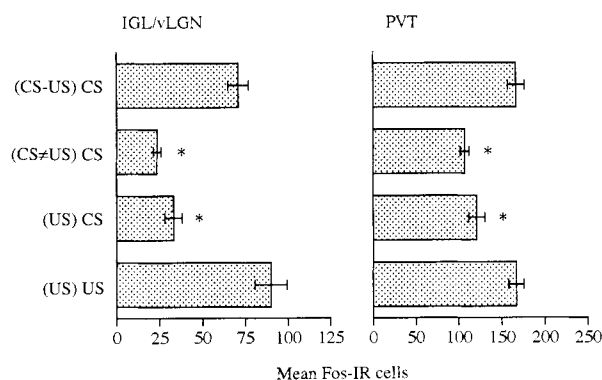


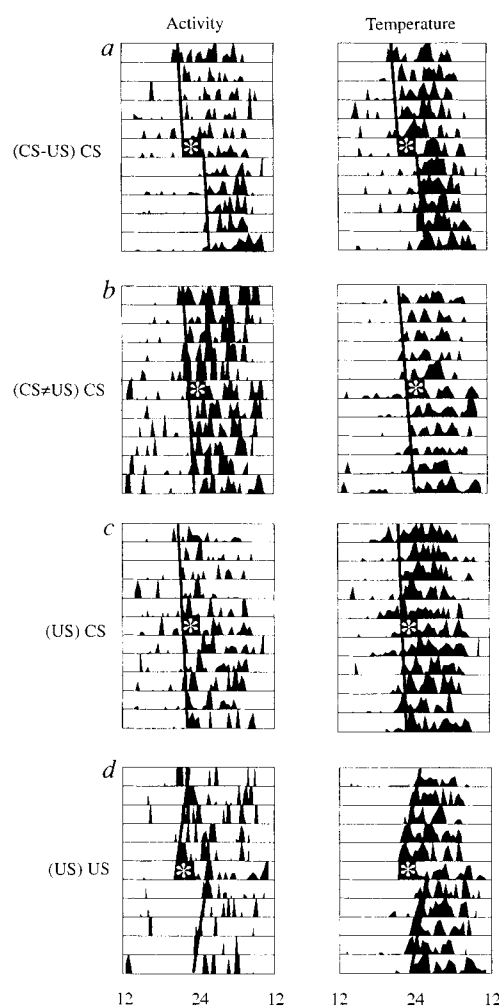
FIG. 3 Means  $\pm$  s.e.m. number of Fos-immunoreactive (Fos-IR) cells in the IGL/vLGN, unilaterally, and PVT calculated for the same animals as in Fig. 2. ANOVA (IGL/vLGN,  $F(3,16) = 38.5$ ,  $P < 0.0001$ ; PVT,  $F(3,16) = 14.03$ ,  $P < 0.0001$ ; asterisk indicates difference from groups (CS-US)CS and (US)US,  $P < 0.05$ .

FIG. 4 Representative records of activity and body temperature from animals kept in constant darkness before and after exposure to the CS in a, group (CS-US)CS; b, group (CS  $\neq$  US)CS; c, group (US)CS; and d, to the US in group (US)US. Each line represents a single 24-h period. Successive days are plotted from top to bottom. The asterisks indicate the time of exposure to the test stimuli. Phase-shifts in the activity rhythm are demonstrated by the presence of a difference between the eye-fitted lines connecting the onset of activity for a period of 5 days before and after exposure to the stimuli. Phase shifts in the temperature rhythm are demonstrated by the presence of a difference between the eye-fitted lines connecting the daily rise in body temperature before and after exposure to the stimuli.

**METHODS.** Rats were implanted intraperitoneally with precalibrated biotelemetry transmitters (Mini Mitter, Sunriver, OR) and housed individually in flat black painted plastic cages ( $36 \times 24 \times 20$  cm). Spontaneous activity and body temperature were recorded at 10-min intervals using the Mini Mitter data acquisition system with Dataquest III software. Activity and body temperature data were smoothed with a 30-min moving average. Phase shifts in activity rhythms were calculated as previously described<sup>29</sup> by subtracting the actual time of activity onset on five days following exposure to the test stimuli from the projected time of activity onset before exposure. Phase shifts in temperature rhythms were calculated by subtracting the time of temperature rise above the mean on five days following exposure to the test stimuli from the projected time of temperature rise before exposure to the test stimuli.

On the sixth day in constant darkness, during the third hour of the subjective night, animals were presented with the test stimuli (CS or US) for 20 min. As illustrated by the examples shown in Fig. 4, presentation of the CS to animals previously given CS-US pairings (group (CS-US)CS,  $n = 3$ ) induced strong phase delays in both activity ( $2.29 \pm 0.19$  h) and temperature ( $2.23 \pm 0.25$  h) rhythms. Similar effects were seen after presentation of the US (group (US)US,  $n = 3$ ; activity,  $2.74 \pm 0.08$  h; temperature,  $2.42 \pm 0.06$  h). Presentation of the CS to animals in group (CS  $\neq$  US)CS ( $n = 3$ ) and group (US)CS ( $n = 3$ ) had no effect on rhythms, indicating that explicit pairing of the CS and US was a necessary condition for the induction of phase shifts by the CS. Thus, the CS that was capable of inducing Fos expression in the SCN was also able to bring about phase shifts in physiological and behavioural rhythms controlled by the SCN pacemaker.

These findings provide evidence that a neutral stimulus, through pairing with light, can acquire the ability to activate a cellular mechanism within the SCN known to mediate photic entrainment. Furthermore, they show that the responsiveness of neural elements within the IGL/vLGN and the PVT that have the potential to integrate photic and non-photoc stimuli and that have direct access to the SCN are likewise altered through conditioning. Thus it is highly likely that the acquired response to the CS seen in the SCN reflects a change in the input from the IGL/vLGN and/or the PVT. Nonetheless, we cannot rule out the involvement of other brainstem and cortical areas as sites for conditioned integration, nor can we exclude the possibility that conditioned



changes could occur in the SCN itself. An identification of these areas as sites for such conditioned stimulus integration could provide a new model for the study of the cellular basis of learning. Finally, it should be stressed that the effects on the circadian system reported here for a non-photoc conditioned stimulus differ in several respects from those of non-photoc stimuli previously reported to cause phase shifts in free running activity rhythms through behavioural activation<sup>27</sup>. Unlike light or the CS in the present study, these non-photoc stimuli are ineffective during the subjective night and do not induce Fos expression in the SCN<sup>23-28</sup>, suggesting that their effect on the pacemaker is mediated by different neural and cellular pathways.

In summary, we find that a stimulus that through associative learning comes to predict light onset can induce those phase-sensitive cellular changes in the SCN and shifts in circadian rhythms that are normally induced by light itself. The robustness of the effect of the conditioned stimulus suggests that the circadian system is amenable to conditioned stimulus control. We propose that similar effects may occur in humans and events that regularly occur in association with the daily rising of the sun or with the onset of artificial light may acquire the ability to reset the circadian pacemaker. □

Received 13 October; accepted 1 December 1995.

1. Morin, L. P. *Brain Res. Rev.* **67**, 102-127 (1994).
2. Pittendrigh, C. S. in *Handbook of Behavioral Neurobiology Biological Rhythms* (ed. Ashchoff, J.) 95-124 (Plenum, New York, 1981).

3. Takahashi, J. S. A. *Rev. Neurosci.* **18**, 531–553 (1995).
4. Kornhauser, J. M., Nelson, D. E., Mayo, K. E. & Takahashi, J. S. *Neuron* **5**, 127–134 (1990).
5. Rusak, B., Robertson, H. A., Wisden, W. & Hunt, S. P. *Science* **248**, 1237–1240 (1990).
6. Schwartz, W. J., Takeuchi, J., Shannon, W., Davis, E. M. & Aronin, N. *Neuroscience* **58**, 573–583 (1994).
7. Amir, S. & Robinson, B. *Neuroscience* **69**, 1005–1011 (1995).
8. Moore, R. Y. & Lenn, N. J. *J. comp. Neurol.* **146**, 1–14 (1972).
9. Johnson, R. F., Moore, R. Y. & Morin, L. P. *Brain Res.* **402**, 301–312 (1988).
10. Card, J. P. & Moore, R. Y. *J. comp. Neurol.* **284**, 135–147 (1989).
11. Moore, R. Y. & Card, J. P. *J. comp. Neurol.* **344**, 403–430 (1994).
12. Rusak, B., Meijer, J. H. & Harrington, M. E. *Brain Res.* **493**, 283–291 (1989).
13. Harrington, M. E. & Rusak, B. *J. biol. Rhythms* **1**, 309–325 (1986).
14. Picard, G. E. *Neurosci. Lett.* **55**, 211–217 (1985).
15. Hickey, T. L. & Spear, P. D. *Exp. Brain Res.* **24**, 523–529 (1976).
16. Shinohara, K., Tominga, K., Fukuhara, C., Otori, Y. & Inouye, S.-I. *T. Neuroscience* **56**, 813–822 (1993).
17. Harrington, M. E., Nance, D. M. & Rusak, B. *Brain Res.* **410**, 275–282 (1987).
18. Card, J. P. & Moore, R. Y. *J. comp. Neurol.* **206**, 390–396 (1982).
19. Edelstein, K. & Amir, S. *Brain Res.* **600**, 254–258 (1995).
20. Davidowa, H. & Albrecht, D. *Behav. Brain Res.* **50**, 127–133 (1992).
21. Moga, M. M., Weis, R. P. & Moore, R. Y. *J. comp. Neurol.* **359**, 221–238 (1995).
22. Mrosovsky, N. *J. comp. Physiol.* **162**, 35–46 (1988).
23. Wickland, C. R. & Turek, F. W. *Am. J. Physiol.* **261**, R1109–R1117 (1991).
24. Van Reeth, O. & Turek, F. W. *Nature* **339**, 49–51 (1989).
25. Zhang, Y., Van Reeth, O., Zee, P. C., Takahashi, J. S. & Turek, F. W. *Neurosci. Lett.* **164**, 203–208 (1993).
26. Janik, D. & Mrosovsky, N. *NeuroReport* **3**, 575–578 (1992).
27. Mead, S. et al. *J. Neurosci.* **12**, 2516–2522 (1992).
28. Cutrera, R. A., Kalsbeek, A. & Pévet, P. *Brain Res.* **602**, 14–20 (1993).
29. Daan, S. & Pittendrigh, C. S. *J. comp. Physiol.* **106**, 253–266 (1976).

ACKNOWLEDGEMENTS. We thank B. Robinson for his expert technical assistance. The work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR, Québec).

## Promiscuous coupling between the sulphonylurea receptor and inwardly rectifying potassium channels

Carina Ämmälä\*, Andrew Moorhouse\*, Fiona Gribble\*, Rebecca Ashfield†, Peter Proks\*, Paul A. Smith\*, Hiroshi Sakura\*, Barbara Coles\*, Stephen J. H. Ashcroft† & Frances M. Ashcroft\*‡

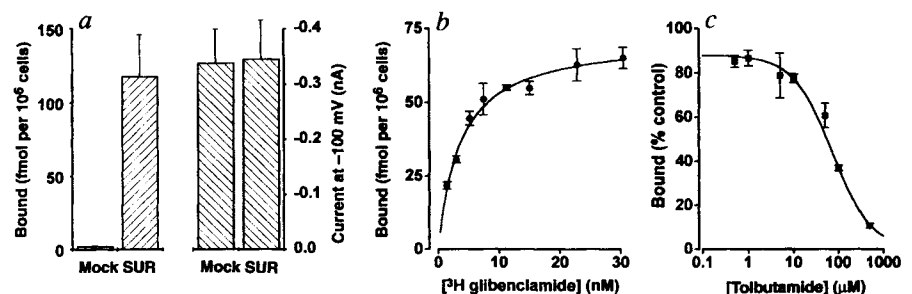
\* University Laboratory of Physiology, Parks Road Oxford, OX1 3PT, UK  
 † Nuffield Department of Clinical Biochemistry, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK

SULPHONYLUREAS are a class of drugs widely used to treat non-insulin-dependent diabetes mellitus. These drugs act by binding to a sulphonylurea receptor (SUR) in the pancreatic  $\beta$ -cell membrane which inhibits an ATP-sensitive potassium (K-ATP) channel and thereby stimulates insulin secretion. There has

been much debate as to whether SUR and the K-ATP channel are the same or separate proteins, whether SUR confers ATP-sensitivity on an ATP-insensitive pore-forming subunit, and whether sulphonylureas can also modulate other types of K-channel. We show here that SUR itself does not possess intrinsic channel activity but that it endows sulphonylurea sensitivity on several types of inwardly-rectifying K-channels. It does not necessarily confer ATP-sensitivity on these channels.

Cloning of a high-affinity sulphonylurea receptor of relative molecular mass 140,000 from insulinoma cells revealed it to be a member of the ATP-binding-cassette (ABC) transporter family, with two nucleotide-binding folds<sup>1</sup>. Mutations in this protein result in persistent hyperinsulinaemic hypoglycaemia of infancy, a disease associated with unregulated insulin secretion<sup>2</sup>. This indicates that SUR acts as a regulator of insulin secretion, possibly by interaction with the  $\beta$ -cell K-ATP channel. Expression of SUR in *Xenopus* oocytes did not result in novel channel activity or confer sulphonylurea sensitivity on the inwardly rectifying ( $K_{ir}$ ) K-channels  $K_{ir}1.1a$ ,  $K_{ir}2.1$  or  $K_{ir}3.4$  (ref. 1). However we were unable to measure [<sup>3</sup>H]glibenclamide binding to *Xenopus* oocytes injected with messenger RNA encoding SUR, indicating that the oocyte is not an adequate expression system for

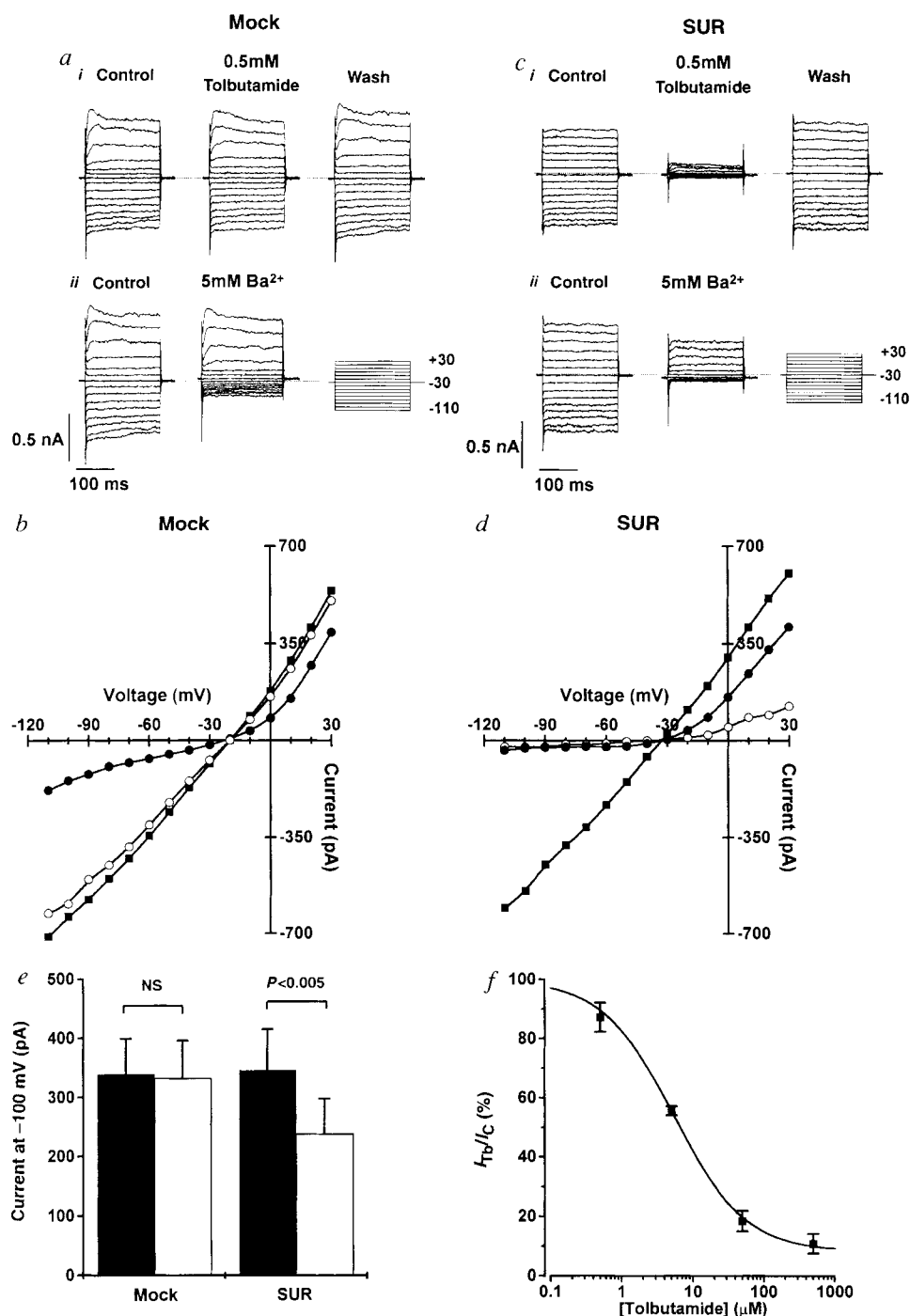
FIG. 1a, Left: [<sup>3</sup>H]glibenclamide binding (fmol per 10<sup>6</sup> cells) to whole HEK293 cells, mock transfected ( $n = 5$ ) or transfected with SUR ( $n = 8$ ), measured 48 h after transfection. Right: Mean whole-cell current amplitude at -100 mV measured in mock-transfected cells ( $n = 33$ ) or cells transfected with SUR ( $n = 24$ ), measured 48 h after transfection. b, [<sup>3</sup>H]glibenclamide binding to whole HEK293 cells transfected with SUR. Specific [<sup>3</sup>H]glibenclamide bound (fmol per 10<sup>6</sup> cells) is plotted against free [<sup>3</sup>H]glibenclamide concentration. Specific binding was determined by subtracting the non-specific binding measured in the presence of 1  $\mu$ M unlabelled glibenclamide. Each point is the mean of 3 replicates. The line is the best fit of a single binding site model to the data.  $K_d = 3.6$  nM and  $B_{max} = 72$  fmol per 10<sup>6</sup> cells. c, Inhibition of [<sup>3</sup>H]glibenclamide binding by tolbutamide measured in the presence of 1.5 nM [<sup>3</sup>H]glibenclamide. The data were fitted to the equation: per cent [<sup>3</sup>H]glibenclamide bound =  $100 / (1 + (K_d/[G])(1 + ([Tb]/K_i)))$ , where  $K_d$  is the dissociation coefficient for [<sup>3</sup>H]glibenclamide binding measured in parallel experiments (3.6 nM),  $K_i$  is the inhibition coefficient for tolbutamide displacement of [<sup>3</sup>H]glibenclamide binding,  $[G]$  is the glibenclamide concentration (1.5 nM),  $[Tb]$  is the tolbutamide concentration. The best fit was obtained with  $K_i = 40$   $\mu$ M. METHODS. The coding sequence of hamster  $\beta$ -cell SUR<sup>1</sup> was subcloned into the vector pcDNA3. HEK293 cells were transiently transfected with this plasmid using lipofectin, as described by the manufacturer. Mock-transfected cells received no DNA but were otherwise treated exactly the same. The transfection efficiency was assessed by  $\beta$ -galactosidase assay. Cells were cultured before and after transfection in MEM medium, supplemented with 10% fetal calf serum, 100 U ml<sup>-1</sup> penicillin and 0.1 mg ml<sup>-1</sup> streptomycin, at 37 °C in an atmosphere of humidified air (95%) and CO<sub>2</sub> (5%). For



measurement of SUR expression, cells ( $\sim 5 \times 10^7$ ) were trypsinized, washed twice with 10 ml of HEPES-Krebs buffer containing (mM): 119 NaCl, 4.75 KCl, 5 NaHCO<sub>3</sub>, 2.54 CaCl<sub>2</sub>, 1.2 MgSO<sub>4</sub>, 1.2 KH<sub>2</sub>PO<sub>4</sub> and 20 HEPES (pH 7.4 with NaOH) and resuspended at a density of  $8.5 \times 10^7$  cells per ml in HEPES-Krebs buffer. Cells were then incubated for 1–2 h at room temperature with 1–30 nM [<sup>3</sup>H]glibenclamide (DuPont NEN; 47 Ci mmol<sup>-1</sup>) at a density of  $1.3$ – $2.0 \times 10^6$  cells per assay tube in 0.4 ml of HEPES-Krebs buffer in the absence or presence of tolbutamide (0.5–500  $\mu$ M). Bound radioactivity was separated from free by rapid filtration on Whatman GF/C filters soaked in HEPES-Krebs buffer; filters were washed four times with 5 ml of buffer at 4 °C. Radioactivity was determined by liquid scintillation spectrometry using an external standard. Non-specific binding was determined in parallel samples which contained 1  $\mu$ M unlabelled glibenclamide. Of the total radioactivity added, 5–10% was bound specifically and approximately 0.4% non-specifically. Methods used to measure whole-cell currents were the same as described in the legend to Fig. 2.

‡ To whom correspondence should be addressed.

FIG. 2 a, Whole-cell currents recorded before, during and after application of 0.5 mM tolbutamide (i), or before and during addition of 5 mM barium (ii), to the external solution, from a mock-transfected HEK293 cell in response to voltage steps from  $-110$  mV to  $+30$  mV from a holding potential of  $-30$  mV (ii, lower right). b, Current-voltage ( $I$ - $V$ ) relationships recorded for the same cell in control solution (■) and in the presence of 5 mM  $\text{BaCl}_2$  (●) or 0.5 mM tolbutamide (○). c, Whole-cell currents recorded before, during and after addition of 0.5 mM tolbutamide (i), or 5 mM barium (ii), to the external solution, from a HEK293 cell transfected with SUR. Same voltage steps as in a. d, Current-voltage ( $I$ - $V$ ) relationships recorded for the same cell in control solution (■) and in the presence of 5 mM  $\text{BaCl}_2$  (●) or 0.5 mM tolbutamide (○). e, Mean current amplitude at  $-100$  mV measured in mock-transfected cells ( $n = 24$ ) or cells transfected with SUR ( $n = 33$ ) in the absence (filled bars) or presence (open bars) of 0.5 mM tolbutamide. f, Mean relationship between tolbutamide concentration and the whole-cell current, expressed as a fraction of its amplitude in the absence of the drug, for 6 cells transfected with SUR. The lines are drawn to a modified form of the Hill equation (see below) with  $K_d = 5.3 \mu\text{M}$ ,  $n = 0.9$ ,  $I_0 = \text{zero}$ . The difference in the  $K_d$  for tolbutamide inhibition of whole-cell currents and the  $K_i$  for displacement of [ $^3\text{H}$ ]glibenclamide binding has been observed before in native membranes<sup>3</sup>. METHODS. Whole-cell currents were studied 48–72 h after transfection, using the standard whole-cell configuration of the patch clamp technique. The pipette solution contained (mM): 107 KCl, 1.2  $\text{MgCl}_2$ , 1  $\text{CaCl}_2$ , 10 EGTA, 5 HEPES (pH 7.2 with KOH; total  $\text{K}^+ \sim 140$  mM) and 0.3 mM ATP. The bathing solution contained (mM): 40 KCl, 100 NaCl, 2.6  $\text{CaCl}_2$ , 1.2  $\text{MgCl}_2$ , 5 HEPES (pH 7.4). Holding potential,  $-30$  mV. Tolbutamide dose-response curves were obtained for each cell using all four drug concentrations. Test solutions were alternated with control solutions and the current ( $I$ ) is plotted as a fraction of the mean ( $I_0$ ) of that obtained in control solution before and after exposure to the drug. Test solutions were applied in random order. Dose-response curves were fitted to the Hill equation:  $I/I_0 = I_0 + ((1 - I_0)/(1 + ([\text{Tb}]/K_d)^n))$  where  $[\text{Tb}]$  is the tolbutamide concentration,  $K_d$  is the tolbutamide concentration at which inhibition is half maximal,  $n$  is the Hill coefficient and  $I_0$  is an offset which



describes the fact that in some cells the maximum block was not 100%. The data are presented as mean  $\pm$  s.e.m. and vertical lines indicate one s.e.m. Statistical significance was tested using the paired  $t$ -test.

SUR and raising the possibility that SUR may in fact act as an ion channel itself or interact with  $\text{K}_{\text{ir}}$  channels.

In contrast to expression in *Xenopus* oocytes, expression of SUR in HEK293 cells produced a dramatic increase in specific binding of [ $^3\text{H}$ ]glibenclamide (Fig. 1a). The mean  $K_d$  for binding to whole cells was  $6.8 \pm 1.8$  nM ( $n = 3$ ; Fig. 1b), comparable to that found for the native  $\beta$ -cell receptor (range 0.3–7 nM; ref. 3) and the mean maximum binding capacity was  $230 \pm 85$  fmol per  $10^6$  cells ( $n = 3$ ) or 138,500 binding sites per cell (assuming all cells are equally well transfected). The affinity for tolbutamide was less, the  $K_i$  for displacement of [ $^3\text{H}$ ]glibenclamide binding being  $59.3 \pm 13.4 \mu\text{M}$ ,  $n = 3$  (Fig. 1c). Similar values have been reported

for native  $\beta$ -cells<sup>3</sup>. Because glibenclamide blocks  $\beta$ -cell K-ATP currents irreversibly<sup>4</sup>, we used tolbutamide to test sulphonylurea sensitivity in electrophysiological experiments.

HEK293 cells express an endogenous  $\text{K}_{\text{ir}}$  current which is observed in both non-transfected and mock-transfected cells (Figs 1a; 2a, b, e). This current was blocked by  $\text{Ba}^{2+}$  but not by 0.5 mM tolbutamide (Fig. 2a), a concentration that maximally inhibits native K-ATP channels<sup>5</sup>. Transfection with SUR did not alter the magnitude of the endogenous whole-cell current: current amplitudes at  $-100$  mV were  $-339 \pm 60$  pA ( $n = 33$ ) in mock-transfected cells compared to  $-346 \pm 69$  pA ( $n = 24$ ) in SUR-transfected cells (Figs 1a, 2e). We therefore conclude that SUR



FIG. 3 *ai*, Whole-cell currents recorded before, during and after addition of 0.5 mM tolbutamide to the external solution, from a HEK293 cell transfected with  $K_{ir}1.1a$  in response to voltage steps from  $-110$  mV to  $+30$  mV from a holding potential of  $-30$  mV. *aii*, Current-voltage ( $I$ - $V$ ) relationships recorded for the same cell in control solution (■) and in the presence of 5 mM  $BaCl_2$  (●) or 0.5 mM tolbutamide (○). *bi*, Whole-cell currents recorded before, during and after addition of 0.5 mM tolbutamide to the external solution, from a HEK293 cell transfected with  $K_{ir}1.1a + SUR$ . Same voltage steps as in *ai*. *bii*,  $I$ - $V$  relationships recorded for the same cell in control solution (■) or in the presence of 5 mM  $BaCl_2$  (●) or 0.5 mM tolbutamide (○). *ci*, Whole-cell currents recorded before, during and after addition of 0.5 mM tolbutamide to the external solution, from a HEK293 cell transfected with  $K_{ir}6.1$ . Same voltage steps as in *ai*. *cii*,  $I$ - $V$  relationships recorded for the same cell in control solution (■) or in the presence of 5 mM  $BaCl_2$  (●) or 0.5 mM tolbutamide (○). *di*, Whole-cell currents recorded before, during and after addition of 0.5 mM tolbutamide to the external solution, from a HEK293 cell transfected with  $K_{ir}6.1 + SUR$ . Same voltage procedure as in *ai*. *dii*,  $I$ - $V$  relationships recorded for the same cell in control solution (■) or in the presence of 5 mM  $BaCl_2$  (●) or 0.5 mM tolbutamide (○). *e*, *f*, Relationship between tolbutamide concentration and the whole-cell current, expressed as a fraction of its amplitude in the absence of the drug, for a cell transfected with  $K_{ir}1.1a + SUR$  (*e*) and for a cell transfected with  $K_{ir}6.1 + SUR$  (*f*). The lines are drawn to the modified Hill equation with  $K_d = 9.9 \mu M$ ,  $n = 1.3$ ,  $I_0 = 0$  and  $K_d = 11.8 \mu M$ ,  $n = 1.3$ ,  $I_0 = 16$ .

**METHODS.** As for Fig. 2. The total DNA concentration was kept constant in cotransfections and the ratio of SUR to  $K_{ir}1.1a$  and  $K_{ir}6.1$  was either 1:1 or 4:1. There were no differences in current amplitude or sulphonylurea sensitivity with these ratios. Cells were assumed to express  $K_{ir}1.1a$  or  $K_{ir}6.1$  if their currents were larger than 1.95 standard deviations above the mean current (95% confidence limits) observed in mock-transfected cells in parallel transfections. Only these cells were used for analysis. Cotransfection with SUR did not alter the current amplitude, which was  $-2446 \pm 667$  ( $n = 9$ ) compared with  $-2034 \pm 186$  pA ( $n = 25$ ) for  $K_{ir}1.1a$  and  $K_{ir}1.1a + SUR$ , respectively, and  $-1,617 \pm 337$  pA ( $n = 10$ ) compared with  $-1,668 \pm 293$  pA ( $n = 16$ ) for  $K_{ir}6.1$  and  $K_{ir}6.1 + SUR$ , respectively. For  $K_{ir}6.1 + SUR$  transfected cells, 0.5 mM tolbutamide produced  $> 10\%$  inhibition in 14/16 cells, compared with 0/7 cells in  $K_{ir}6.1$  transfected cells. For  $K_{ir}1.1a + SUR$  transfected cells, 0.5 mM tolbutamide produced  $> 10\%$  inhibition in 7/25 cells, compared with 0/9 cells in  $K_{ir}1.1a$  transfected cells.  $K_{ir}6.1$  is not endogenously expressed in HEK293 cells<sup>9</sup>.

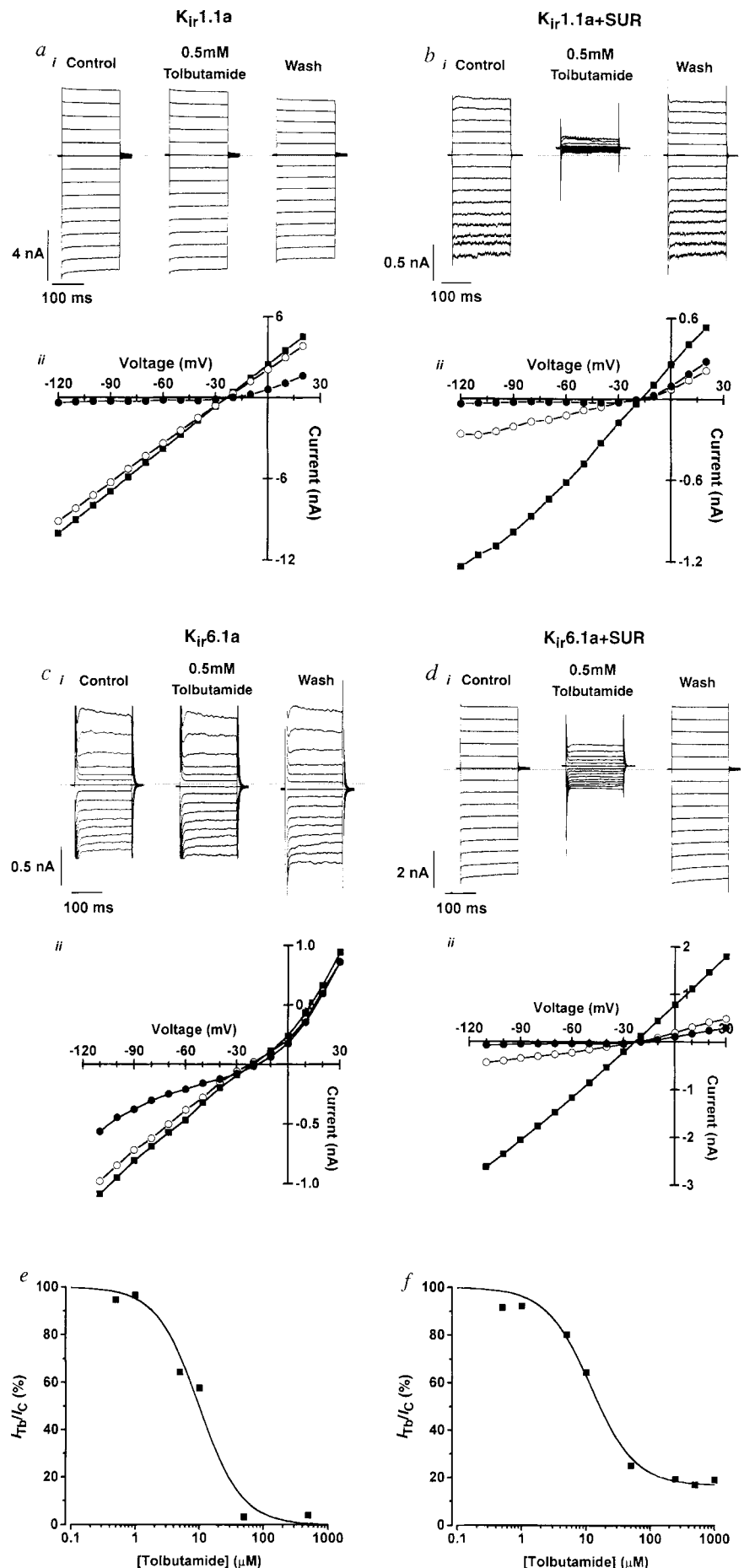


TABLE 1 SUR endows several types of  $K_v$  channels with similar sensitivity to tolbutamide

|                  | IpA (–100 mV)     | $K_d$ ( $\mu$ M)  | Tolbutamide inhibition |  | n |
|------------------|-------------------|-------------------|------------------------|--|---|
|                  |                   |                   | Hill coefficient       |  |   |
| SUR (0.3 mM ATP) | –432 $\pm$ 126    | 5.4 $\pm$ 0.2     | 0.85 $\pm$ 0.09        |  | 7 |
| SUR (5 mM ATP)   | –803 $\pm$ 184 NS | 10.4 $\pm$ 3.3 NS | 0.77 $\pm$ 0.08 NS     |  | 6 |
| SUR+ $K_v$ 1.1a  | –1275 $\pm$ 255   | 22.6 $\pm$ 7.1 NS | 0.83 $\pm$ 0.15 NS     |  | 5 |
| SUR + $K_v$ 6.1  | –2442 $\pm$ 506   | 17.7 $\pm$ 4.1*   | 0.75 $\pm$ 0.15 NS     |  | 5 |

Mean values for cells transfected with SUR and dialysed with either 0.3 mM or 5 mM ATP and for cells transfected with  $K_v$ 1.1a + SUR or  $K_v$ 6.1 + SUR (both dialysed with 0.3 mM ATP).

\*  $P < 0.05$  against SUR (0.3 mM ATP). NS, not significant. A one-way ANOVA was to validate the use of Student's *t*-test. Significance was tested using Student's *t*-test assuming unequal variance.

does not itself form an ion channel or increase expression of endogenous  $K_v$  channels. In 13 of 24 (64%) cells transfected with SUR, however, tolbutamide inhibited the endogenous current by more than 10% (Fig. 2c–e; mean inhibition, 33  $\pm$  7% ( $n = 24$ ) compared with 6  $\pm$  2% ( $n = 33$ ) in mock-transfected cells). It is unlikely that all cells studied were transfected with SUR as our transfection efficiency was roughly 70%; furthermore, even in transfected cells, not all  $K_v$  channels may be coupled to SUR. This may explain why 0.5 mM tolbutamide did not produce complete inhibition of inward currents in all cells tested. The mean dose–response curve for tolbutamide inhibition was best fitted with a  $K_d$  of 5.3  $\mu$ M and a Hill coefficient of 0.9 (Fig. 2f, Table 1), values which are similar to those found for K-ATP channel inhibition in  $\beta$ -cells<sup>6</sup>. This result indicates that SUR can confer tolbutamide sensitivity on the endogenous  $K_v$  current. An endogenous voltage-gated outward K-current was also observed in a few cells: tolbutamide was without effect on this current even in those SUR-transfected cells in which the drug blocked the  $K_v$  currents.

It has been suggested that SUR might endow ATP-sensitivity on the  $\beta$ -cell K-ATP channel by virtue of its nucleotide-binding domains<sup>1</sup>. This does not seem to be the case for the endogenous  $K_v$  current as mean currents at –100 mV were unaffected by inclusion of 5 mM ATP in the pipette solution (Table 1) and we observed no increase in current during the first 5 min after establishment of the whole-cell configuration, as would be expected for an ATP-sensitive current when ATP is dialysed from the cell<sup>5,14</sup>. The  $K_d$  for tolbutamide inhibition was also unaffected (Table 1). Increasing the intracellular concentration of ATP from 0.3 to 5 mM thus does not seem to affect the coupling between SUR and the endogenous  $K_v$  channel in HEK293 cells.

The lack of ATP-sensitivity of the endogenous  $K_v$  current in HEK293 cells indicates that it is not a K-ATP channel and further suggests that SUR may couple to additional types of  $K_v$  channel (see ref. 7 for nomenclature). We therefore examined whether SUR couples to two other members of the  $K_v$  family,  $K_v$ 1.1a (ROMK1) which is primarily expressed in the kidney<sup>8</sup> and  $K_v$ 6.1 (uKATP1) which is a ubiquitously expressed ATP-sensitive K channel<sup>9</sup>. Figure 3 shows that cells transfected with  $K_v$ 1.1a (a) or  $K_v$ 6.1 (c) expressed  $Ba^{2+}$ -sensitive currents which were much

larger than those seen in mock-transfected cells. These currents were insensitive to tolbutamide, the mean current amplitude at –100 mV in the presence of 0.5 mM tolbutamide being 99  $\pm$  0.8% ( $n = 9$ ) and 102  $\pm$  4% ( $n = 7$ ) of that recorded in the absence of the drug, for  $K_v$ 1.1a and  $K_v$ 6.1, respectively. Cotransfection with SUR, however, conferred tolbutamide sensitivity on both  $K_v$ 1.1a and  $K_v$ 6.1 currents (Fig. 3b, d–f and Table 1) but did not alter the current amplitude. In addition, 10  $\mu$ M glibenclamide irreversibly blocked whole-cell currents in cells expressing  $K_v$ 1.1a and SUR.

In kidney epithelial cell membranes, the native K-ATP channel, believed to correspond to  $K_v$ 1.1a, is less sensitive to tolbutamide or glibenclamide<sup>10</sup> than when  $K_v$ 1.1a is coupled to SUR. This suggests that in native membranes  $K_v$ 1.1a may not couple to SUR but to another channel regulator which endows sulphonylurea sensitivity. One possible candidate is the cystic fibrosis gene product CFTR, another member of the ABC transporter family, which acts both as an ion channel and as a channel regulator<sup>11</sup> and which is modulated by sulphonylureas<sup>12</sup>.

We conclude that SUR does not act as an ion channel itself when expressed in HEK293 cells. Instead, it seems to act as a regulator of  $K_v$  channels, endowing them with sulphonylurea sensitivity. Our data indicate that SUR is promiscuous and can couple to more than one type of  $K_v$  channel, as SUR regulates both  $K_v$ 1.1a and  $K_v$ 6.1, and the endogenous  $K_v$  channel of HEK293 cells: in addition, SUR must also regulate the native  $\beta$ -cell K-ATP channel<sup>6</sup>. The results suggest the  $K_d$  for tolbutamide inhibition is determined primarily by SUR and is little influenced by the  $K_v$  channel with which it interacts. Our data also demonstrate that SUR does not confer ATP-sensitivity on the endogenous  $K_v$  channel of HEK293 cells, although we cannot exclude the possibility that it regulates the ATP-sensitivity of the  $\beta$ -cell K-ATP channel. Finally, our observation that SUR couples to several types of  $K_v$  channel means that the presence of sulphonylurea sensitivity cannot be taken to indicate the presence of K-ATP channels, as has been widely assumed in the literature.

**Note added in proof:** The  $\beta$ -cell K-ATP channel has now been shown to be a complex of SUR and  $K_v$ 6.2<sup>13,14</sup>. As  $K_v$ 6.2 does not express independently it remains unclear whether SUR confers ATP sensitivity on the K-ATP channel. □

Received 9 November; accepted 29 December 1995.

1. Aguilar-Bryan, L. et al. *Science* **268**, 423–425 (1995).
2. Thomas, P. M. et al. *Science* **268**, 425–429 (1995).
3. Ashcroft, S. J. H. & Ashcroft, F. M. *Biochim biophys Acta* **1175**, 45–59 (1992).
4. Zünckler, B. J., Lenzen, S., Männer, K., Panten, U. & Trube, G. *Nannyn-Schmiedeberg's Archs Pharmacol.* **337**, 225–230 (1988).
5. Trube, G., Rorsman, P. & Ohno-Shosaku, T. *Pflügers Arch.* **407**, 493–499 (1986).
6. Ashcroft, S. J. H. & Ashcroft, F. M. *Cell Sig.* **2**, 197–214 (1990).
7. Doupnik, C. A., Davidson, N. & Lester, H. A. *Curr. Opin. Cell Biol.* **5**, 268–278 (1995).
8. Ho, K. et al. *Nature* **362**, 31–39 (1993).
9. Inagaki, N. et al. *J. Biol. Chem.* **270**, 5691–5694 (1995).

10. Tsuchiya, T., Wang, W., Geibisch, G. & Welling, P. A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6418–6422 (1992).
11. Higgins, C. F. *Cell* **82**, 693–696 (1995).
12. Sheppard, D. N. & Welsh, M. J. *J. gen. Physiol.* **100**, 573–593 (1992).
13. Inagaki, N. et al. *Science* **270**, 1166–1170 (1995).
14. Sakura, H., Ammälä, C., Smith, P. A., Gribble, F. M. & Ashcroft, F. M. *FEBS Lett.* **377**, 338–344 (1995).

**ACKNOWLEDGEMENTS.** We thank the Wellcome Trust and the British Diabetic Association for support. F.G. was an MRC Clinical Training Fellow. We thank L. Aguilar-Bryan and J. Bryan for the gift of SUR, S. Hebert for  $K_v$ 1.1a and S. Seino for  $K_v$ 6.1.

# Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry

David A. Leopold & Nikos K. Logothetis\*

Division of Neuroscience, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

WHEN the two eyes view dissimilar images, we experience binocular rivalry, in which one eye's view dominates for several seconds and is then replaced by that of the other eye<sup>1,2</sup>. What causes these perceptual changes in the absence of any change in the stimulus? We showed previously that some neurons in monkey cortical area MT show changes in activity during motion rivalry that reflect the perceived direction of motion<sup>3</sup>. To determine whether perception-related modulation of activity occurs in other visual cortical areas, we recorded from individual neurons in V1, V2 and V4 while monkeys reported the perceived orientation of rival gratings of two orthogonal orientations. Many cells, particularly in V4, showed patterns of activity that correlated with the perceptual dominance and suppression of one stimulus. The majority were orientation-selective and could be driven equally well from either eye. It has been previously suggested that binocular rivalry involves reciprocal inhibition between monocular neurons within V1 (for example, see ref. 4), but our results do not support this view; rather, we propose that binocular rivalry arises through interactions between binocular neurons at several levels in the visual pathways, and that similar mechanisms may underlie other multistable perceptual states that occur when viewing ambiguous images.

Two monkeys (*Macaca mulatta*) were trained to fixate on a light spot and to report their perceived orientation of a grating stimulus by pressing levers. Once they had learned to respond rapidly and accurately to orientation changes, periods of rivalrous stimulation (4–12 seconds) were randomly intermixed with periods of congruent stimulation in observation periods of up to 25 seconds.

It was important to ensure that the monkeys were reporting their perceptions reliably rather than pressing levers at random, and we confirmed this in three ways. First, we introduced two types of catch trials in which the orientation of one grating was smoothly replaced after a lever response to yield a coherent binocular stimulus. In some cases the orientation was the same as that indicated by the monkey's last report, and the monkey was expected not to respond, and in other cases the orthogonal grating was faded in, and the monkey was expected to respond immediately to the change. Performance in both the non-rivalrous periods and the catch trials was consistently above 95%. As a second test, we compared the distribution of the durations of rivalry phases with parallel data from human subjects. As average phase durations depend on stimulus parameters that were adjusted for each recording session (such as contrast, size and spatial frequency; see below)<sup>5</sup>, phase durations were normalized to the mean for that session. Figure 1a shows the distributions obtained from monkeys and humans, together with the  $\gamma$ -density functions typically used to describe such distributions. The similarity between the two distributions argues against the possibility of random reporting. As a final control, we varied the contrast for each eye independently. Increasing the stimulus contrast in one eye is known to decrease the time for which that stimulus remains suppressed, while minimally affecting the time for which it remains dominant<sup>5</sup>. Again, a similar relationship between stimulus contrast and eye dominance was seen in monkeys and humans (Fig. 1b). Such consistency is very unlikely to arise by chance.

Having established that monkeys were reporting their perceptions accurately, we recorded from individual neurons in the

foveal V1/V2 (which cannot be clearly distinguished without histology) and in V4, while monkeys reported the orientations of congruent or rivalrous stimuli. Many cells in both V1/V2 and V4 respond best to gratings of a specific orientation, and they also vary in the extent to which they are driven by one or both eyes. Because cell responses during rivalry can best be studied by presenting the neurons with their optimal and non-optimal stimuli simultaneously, we were interested first to accurately determine each neuron's orientation preference. Moreover, as it has been suggested that binocular rivalry is a manifestation of reciprocal inhibition between the two monocular inputs<sup>4</sup>, we were also interested in examining the ocular dominance of each cell. We therefore characterized the receptive field properties of each neuron under non-rivalry conditions, during passive fixation, before testing its response during rivalry. Each neuron was then tested with rivalrous stimulation, with its optimal orientation presented to one eye at random while the orthogonal orientation was presented to the other eye. In total, we characterized 101 neurons from two monkeys (three hemispheres). Figures 1–4 show data from one monkey; Fig. 5 shows pooled data from both monkeys.

Figure 2 shows observation periods for a single V4 neuron with a foveal receptive field, which was binocular and tuned for a right-tilting 45° orientation. The figure shows the spike trains, spike density functions, eye position, and reports of perceived grating orientation (the latter indicated by the red/green bar, in which each colour change represents a lever press). Note that the firing rate is elevated before the monkey reports perceiving the cell's preferred orientation, and often remains high for about the same time as the dominance phase of that orientation. This occurs regardless of whether the preferred orientation is presented to the right or the left eye.

To analyse further the neural activity associated with a perceptual change in the absence of a stimulus change, we constructed peristimulus time histograms (PSTHs) aligned to the animal's reports. Figure 3a shows data for the cell shown in Fig. 2. The activity from periods when the direction of gaze was outside the target was excluded from the analysis. The left panel shows reported transitions to the preferred orientation (green to red in Fig. 2); the right panel shows transitions to the suppression of this orientation. Note the increase in activity before the monkey reports perceiving the cell's optimal orientation, and the decrease before the report of the orthogonal orientation. Similar patterns of activity were seen in V1/V2 neurons (Fig. 4), but another type of behaviour was also seen in V4, in which some neurons discharged in association with the perceptual suppression (rather than activation) of the orientation that was preferred in non-rivalrous stimulation (Fig. 3b).

A quantitative analysis of all the responses revealed several classes of cells that differ in their responses to congruent and rivalrous stimuli (Fig. 5). About one third of the neurons studied were found to modulate their activity during rivalry in accordance with the monkey's report, whereas the rest were either inhibited throughout the rivalrous period or remained unaffected, spiking at the same rate under monocular, binocular or rivalrous conditions. In V1/V2, 6/33 neurons examined showed modulation, of which three were orientation-selective during congruent stimulation, whereas the other three showed no statistically significant preference for either orientation. In V4, 26/68 cells were modulating; twelve of these fired best when their preferred orientation was perceived, six fired when their preferred orientation was suppressed, and the remaining eight showed no preferred orientation during congruent stimulation. All but one of the cells showing response modulation during rivalry were binocular.

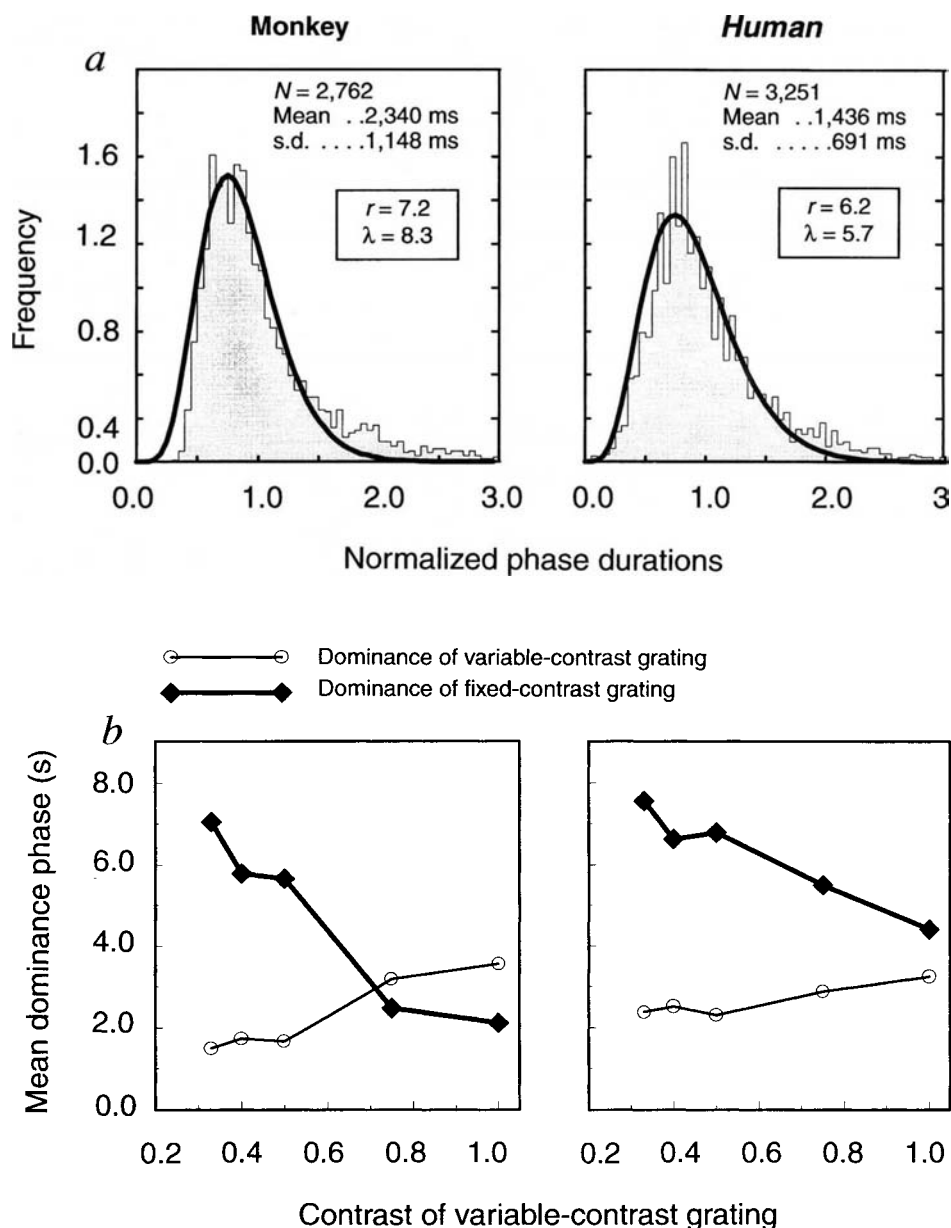
In a previous study on binocular motion rivalry<sup>3</sup>, we showed that some neurons in MT respond when their preferred stimulus is perceived, some respond when it is suppressed, and others are relatively untuned during coherent stimulation but show enhanced selectivity in response to rivalrous stimuli. In addition,

\* To whom correspondence should be addressed.

FIG. 1 Temporal dynamics of binocular rivalry (BR) in monkeys and humans. *a*, The frequency histograms show the distribution of relative phase durations, that is, phase durations expressed as a fraction of the mean phase duration. The smooth, thick, black lines illustrate the approximation of the data with a  $\gamma$  function  $f(x) = \lambda^r / \Gamma(r) x^{r-1} \exp(-\lambda x)$ , where  $\Gamma(r) = (r-1)!$ . The parameters of the theoretical distribution describing the monkey's data were not significantly different ( $r = 5.57$ ,  $\sigma_r = 3.83$ ,  $t_r = 0.41$ , and  $\lambda = 6.02$ ,  $\sigma_\lambda = 4.55$ ,  $t_\lambda = 0.48$ ; two-tailed  $t$ -test) from those obtained in human experiments in this and other laboratories<sup>5,10,11</sup>.

*b*, Effects of variation of stimulus strength on the mean dominance phase. Fixed contrast was set to 1.0. On the abscissa is plotted the varied contrast of the stimulus in one eye, and on the ordinate the mean dominance duration of that eye (circles) and of the eye receiving the grating with the fixed contrast (diamonds). All data were collected from the same animal used to study the cell responses illustrated in Figs 2–4. Frequency histograms were built from the rivalry reports during the physiology sessions. The effects of stimulus strength, which require varying the stimulus contrast, were studied during dedicated psychophysical sessions.

**METHODS.** The monkeys underwent an aseptic surgery, using isoflurane anaesthesia (1.2%–1.5%), for the placement of the head restraint post and the scleral search eye coil. Throughout the surgical procedure, the heart rate, blood pressure and respiration were monitored constantly and recorded every 15 min. Body temperature was kept at 37 °C. Postoperatively, the monkeys were administered an opioid analgesic (buprenorphine hydrochloride 0.02 mg kg<sup>-1</sup>; IM) every 6 h for 2 d, and tylenol (10 mg kg<sup>-1</sup>) and antibiotics (tribissen, 30 mg kg<sup>-1</sup> for 3–5 days. At the end of the training period, another sterile surgical operation was performed to implant a chamber for the electrophysiological recordings. The visual stimuli were generated with an image processing system (MV200 Datacube, Inc.), and were presented on a display monitor (BARCO CDID 7651) placed 97 cm away from the animal. Stereoscopic presentations were accomplished using a liquid crystal polarizer (Tektronix SGS 610), which allowed alternate transmission of images with circularly opposite polarization at the rate of 120 frames s<sup>-1</sup> (60 frames s<sup>-1</sup> for each eye). Humans and monkeys wore glasses that allowed the passage of only every other image to each eye. The monkeys were trained to discriminate the orientation of the grating stimulus. Initially, they were rewarded for each correct response. When they consistently attained better than 95% accuracy, the training continued by reinforcing them only at the end of each observation period; however, feedback as to the correctness of the response was always given by aborting the observation period each time the monkey broke fixation or responded incorrectly. Once they had learned the orientation discrimination task, periods of rivalrous stimulation were randomly intermixed with periods of congruent stimulation. Standard techniques were used for recording from single units<sup>12</sup>. The receptive field (RF) of each isolated neuron was first plotted with a computer controlled bar stimulus. The width and height of the optimally oriented bar were used to determine the orientation, spatial frequency and size of a test-grating. Quantitative tests of each neuron's specificity for orientation, disparity and ocular dominance were then conducted while the monkey performed a fixation task. The average RF size of the V1/V2 and the V4 neurons was found to be 0.52 ± 0.16 and 0.66 ± 0.14 degrees respectively. The



average position, in terms of azimuth and elevation, of the RF's centre was (−0.01, 0.10) degrees for V1/V2, and (0.02, −0.02) degrees for V4. The size of the stimuli, which were centred in the receptive field, was kept as small as possible (usually about 0.7 to 1.2 degrees) to minimize piecemeal rivalry; that is, the perception of a mosaic-like collage consisting of different portions of each eye's stimulus-pattern. Decreasing stimulus size increases the exclusive visibility of a stimulus, and therefore optimizes the conditions for studying correlations between cell and monkey responses. Following the quantitative tests, each neuron was tested during the orientation discrimination task with congruent and rivalrous stimuli. The relationship between eye and preferred orientation was pseudorandomized across rivalrous observation periods, so that in each session an equal number of preferred and non-preferred orientations were presented to each eye. Cells were collected from the foveal representation of the areas V1 and V2. As no histology is yet available, the cells in the two early areas are referred to as V1/V2 neurons. However, based on our preliminary topographical mapping, and on ocular dominance preferences in single penetrations, about half of the tested V1/V2 cells are most likely in striate cortex and the other half in area V2. On the basis of stereotaxic coordinates, RF size and position, and RF properties, most neurons on the prelunate gyrus or the anterior bank of the lunette were in V4.



many neurons continue to respond to their preferred stimulus regardless of the perceptual state. The present study shows that a similar range of response patterns exists in V4 and V1/V2, although areas V1/V2 did not have neurons responding during the suppression of their preferred stimulus. As the duration of dominance and suppression can vary independently as the stimulus strength is changed, it is tempting to speculate that the neurons firing during dominance mediate the perception of one stimulus, whereas those firing during suppression underlie the phenomenal suppression of the other. Interestingly, a number of modulating neurons were influenced by the perceptual requirements of the task. Thus some cells were tuned for orientation (this work) or direction<sup>3</sup> during the discrimination task but showed no preference during the fixation task, whereas others were tuned during rivalrous but not congruent stimulation. Such cells may be med-

iating the effects of selective attention on perception. It remains a puzzle, however, why the discharges of the large number of neurons that continue to fire during suppression is not sufficient to mediate perception. It will be of great interest to determine whether the different response properties are in any way correlated with the anatomical location and connectivity patterns. Preliminary results suggest that many of the modulating neurons in MT are located in deep cortical layers (N.K.L., unpublished results).

It has been proposed that the phenomenal suppression during binocular rivalry is the result of a blockade of monocular information triggered by reciprocal inhibition between monocular neurons in striate cortex<sup>4</sup>. Such a blockade would result in the removal of the representation of the suppressed stimulus at all subsequent sites in the visual pathway. Our results do not support

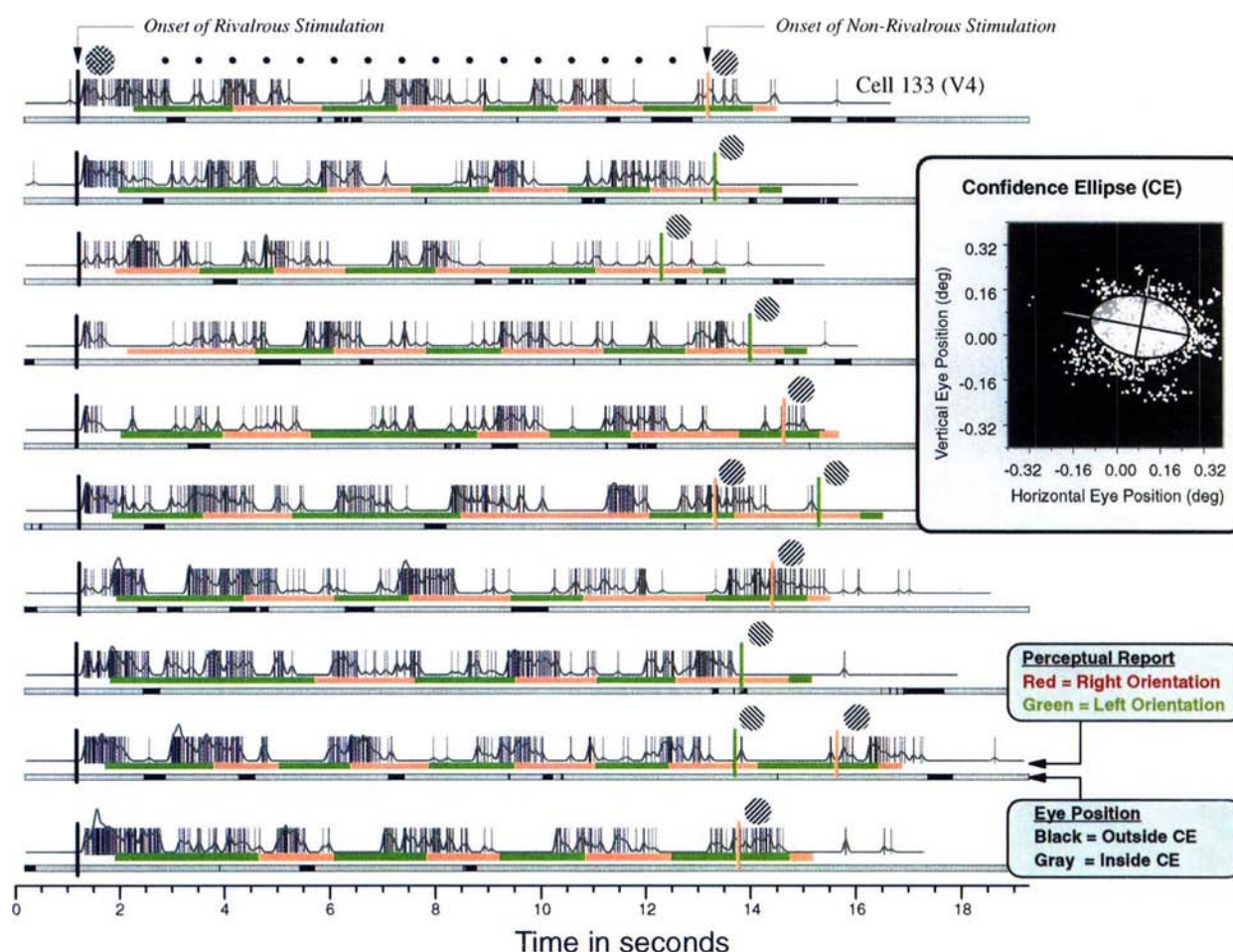


FIG. 2 Single observation periods showing modulations in the firing rate of a V4 neuron as the monkey reports perceptual changes during continued viewing of the same stimulus. Each thin dark-grey vertical line signifies the occurrence of an action potential. Superimposed on the action potentials are spike density functions illustrating the probability of occurrence of a single spike within successive 5-ms periods. The horizontal green and red bars show the alternating phases of perceptual dominance of the left- and right-tilted grating. The black regions on the grey bar show periods in which the eye position was outside the confidence ellipse shown in the inset. This ellipse surrounds the region within which eye positions kept the stimulus within the central part of the neuron's receptive field. This region was computed as follows: mean eye position ( $X, Y$ ) and spike rate  $R$  were first calculated over 35 small time windows (a 250-ms window that was shifted by 50 ms, beginning 1,000 ms before and ending 750 ms after each lever

press) during non-rivalrous periods in which the neuron was stimulated with its preferred orientation. The average ( $\bar{X}, \bar{Y}$ ) of the  $X, Y$  values for which  $R > (R + \sigma_R)$ , was taken as the centre of the ellipse that represents the eye position that 'centres' the stimulus within the RF. The slope and length of the axes of the confidence ellipses correspond to the eigenvectors and eigenvalues of the distribution of the selected eye positions. Only rivalrous trials during which the eye position remained within the confidence ellipse were used to build the PSTHs shown in the next figures. Note that the firing rate of the neuron is elevated immediately before each report of perceiving the cell's preferred orientation, and remains high often for a time-period roughly corresponding to the following dominance phase of the preferred stimulus. In about half of the illustrated observation periods the left-tilted grating is presented to the left and the other half to the right eye.

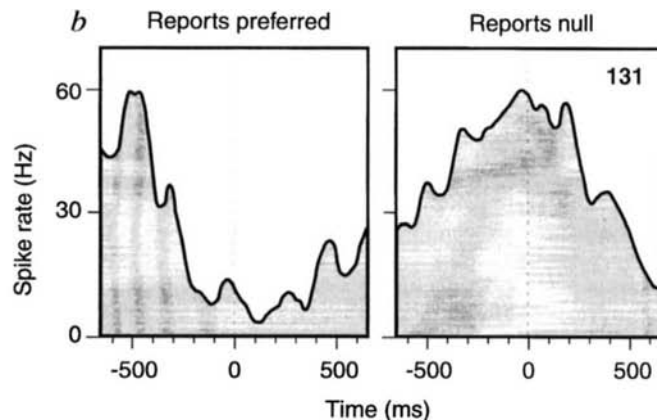
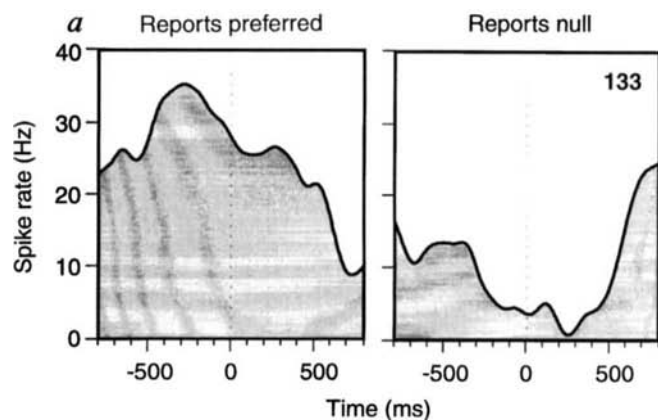


FIG. 3 Peristimulus time histograms for two V4 neurons during BR. The ordinate represents the spike rate and the abscissa the time before and after the monkey's report, depicted by a slashed line. The left and right plots show the activity of the cell for those trials in which the monkey reported the perceptual dominance of the neuron's preferred and non-preferred orientation respectively. Because the phase durations vary greatly within a session, averaging within any time window may often include cell responses from the previous or next report. Thus the selection of a time window—here  $[-650, 650]$  ms—around the animal's response

is arbitrary, usually based on the mean phase duration of the session. *a*, Mean response of the neuron shown in Fig. 2 before and during the report of phenomenal dominance and suppression of a right-tilted grating. *b*, An example of a neuron firing exclusively before and during the suppression of its preferred orientation. The neuron's response is reduced before the monkey reports seeing the cell's optimal orientation, and remains very low for more than 600–700 ms. In contrast, the response gradually increases before the monkey's report that the optimal orientation is suppressed, and remains high for a similar time period (600–700 ms).

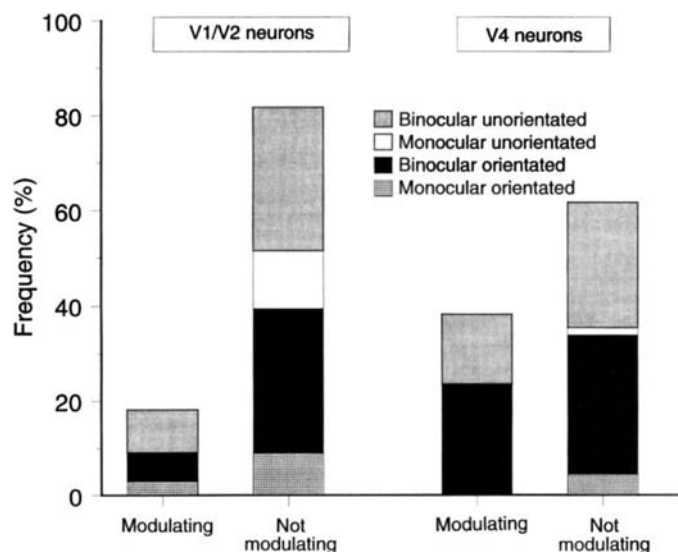


FIG. 5 Distribution of different cell properties within the population of response-modulating and non-modulating cells in the areas V1/V2 and V4. Cells were classified as monocular/binocular, orientated/unorientated and modulating/non-modulating (see text).

this view. First, almost all recorded monocular cells are unaffected by suppression. Second, the frequency of modulating cells is significantly higher in V4 (38%) and MT (43%; ref. 3) than it is in V1/V2 (18%) where the inputs from the two eyes are first combined. Third, the majority of cells in all areas continue to respond to their preferred stimulus even when it is perceptually suppressed; a fact that is consistent with a number of studies showing that suppression of an adapting stimulus during binocular rivalry does not curtail the strength of after-effects<sup>6</sup>. Thus, competition between monocular neurons is unlikely to explain rivalry; instead, we suspect that the rivalry is between alternative stimulus representations that are encoded in the activity of many neurons in different visual areas (among them V4 and MT) implicated in

shape encoding or figure-ground segregation. This hypothesis is supported by a number of psychophysical observations. For example, rivalry can occur even when the conflicting patterns are presented to the same eye, ('monocular rivalry' or 'pattern rivalry')<sup>7</sup>. Also, when two pairs of rivalrous stimuli are presented together, the perceptions are more likely to change in synchrony if members of the two pairs form a coherent figure, even if the components of the figure are presented to opposite eyes<sup>8</sup>.

Finally, the temporal dynamics of rivalry are similar to those experienced when viewing ambiguous figures, such as the Necker cube and other depth reversals<sup>9</sup>. It is therefore possible that a common neural mechanism underlies all these forms of multi-stable perception, and that understanding this mechanism will

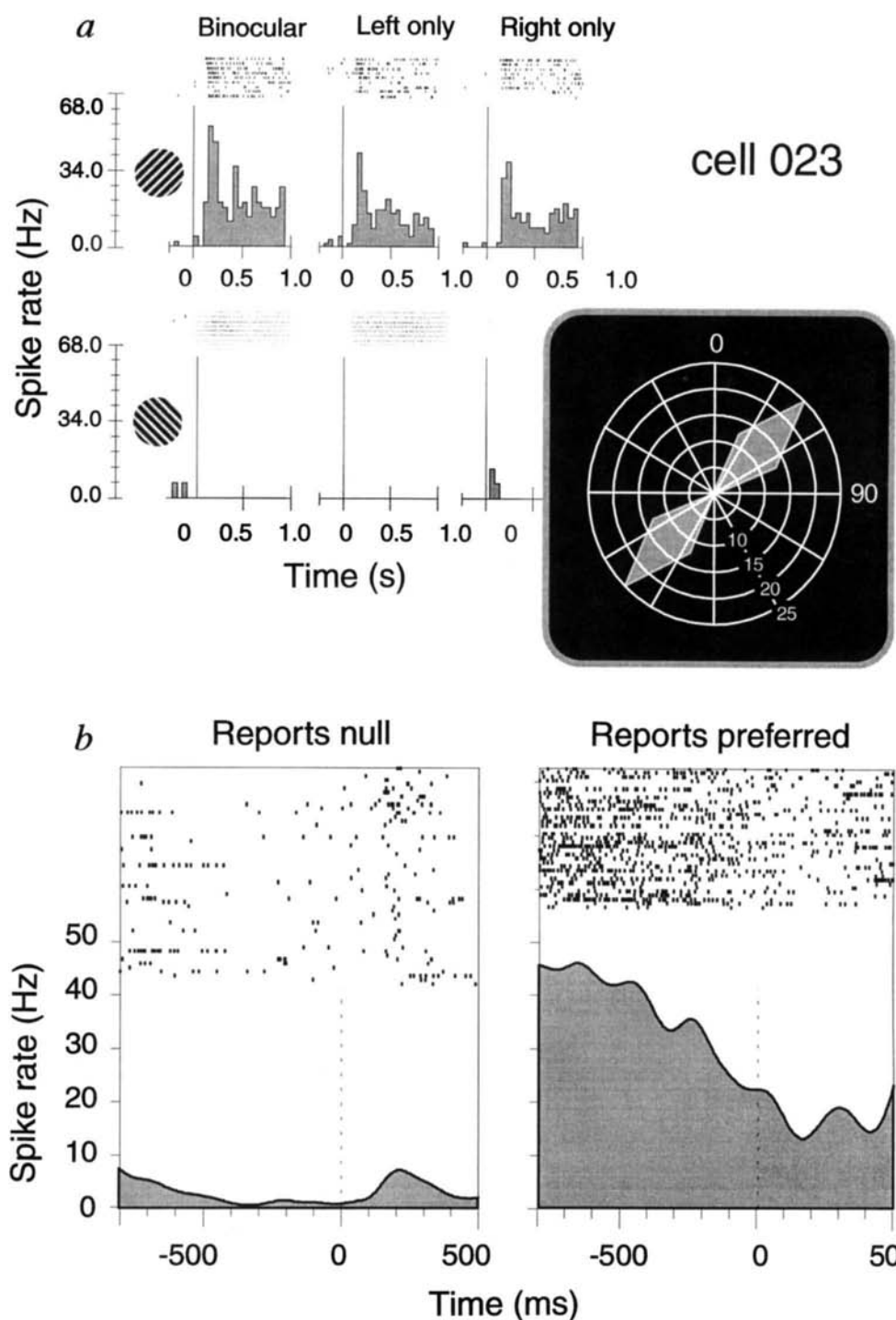


FIG. 4 Response of a V1/V2 neuron to rivalrous stimuli. *a*, PSTHs showing the cell's response to a grating of optimal (above) and of orthogonal-to-optimal (below) orientation, for binocular and monocular presentations. The polar plot shows the neuron's orientation tuning. *b*, Rasters and PSTHs during rivalrous stimulation. Responses are averaged for a time window of  $[-800, 500]$  ms around the monkey's response. Conventions as in Fig. 3.

lead to new insights regarding the neural basis of perceptual organization and of visual awareness. □

Received 17 July; accepted 13 December 1995.

1. Helmholtz, H. V. *Physiological Optics*, translated for Opt. Soc. Am. (1924) from *Handbuch der Physiologischen Optik* 3rd edn (Voss, Hamburg, 1909) (Dover, New York, 1962).
2. Breese, B. *Psychol. Rev.* **3**, 1–65 (1899).
3. Logothetis, N. & Schall, J. *Science* **245**, 761–763 (1989).
4. Blake, R. *Psychol. Rev.* **96**, 145–167 (1989).
5. Levelt, W. *On Binocular Rivalry* (Royal VanGorcum, Assen, 1965).

6. Blake, R. & Rox, R. *Nature* **249**, 488–490 (1974).
7. Campbell, F. & Howell, E. J. *Physiol., Lond.* **225**, 19P–21P (1972).
8. Whittle, P., Bloor, D. & Pocock, S. *Percept. Psychophys.* **4**, 183–188 (1968).
9. Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S. & Bartolini, B. *Kybernetik* **10**, 139–144 (1972).
10. Fox, R. & Herrmann, J. *Percept. Psychophys.* **2**, 432–436 (1967).
11. Walker, P. *Percept. Psychophys.* **18**, 467–473 (1975).
12. Logothetis, N., Pauls, J. & Poggio, T. *Curr. Biol.* **5**, 552–563 (1995).

ACKNOWLEDGEMENTS. We thank F. Crick, J. Pauls, D. Sheinberg, J. Maunsell and J. Assad for comments on the manuscript and D. Murray for technical assistance. This research was supported by an NIH grant to N.K.L., and in part by an AASERT grant from the Office of Naval Research.

# Bax-independent inhibition of apoptosis by Bcl-x<sub>L</sub>

Emily H.-Y. Cheng\*, Beth Levine†, Lawrence H. Boise‡, Craig B. Thompson‡ & J. Marie Hardwick\*§

\* Neurovirology Laboratories, Department of Neurology, Department of Pharmacology and Molecular Sciences, Meyer 6-181, Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, Maryland 21287-7681, USA

† Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

‡ Howard Hughes Medical Institute and Departments of Medicine, Molecular Genetics and Cell Biology, University of Chicago, Chicago, Illinois 60637, USA

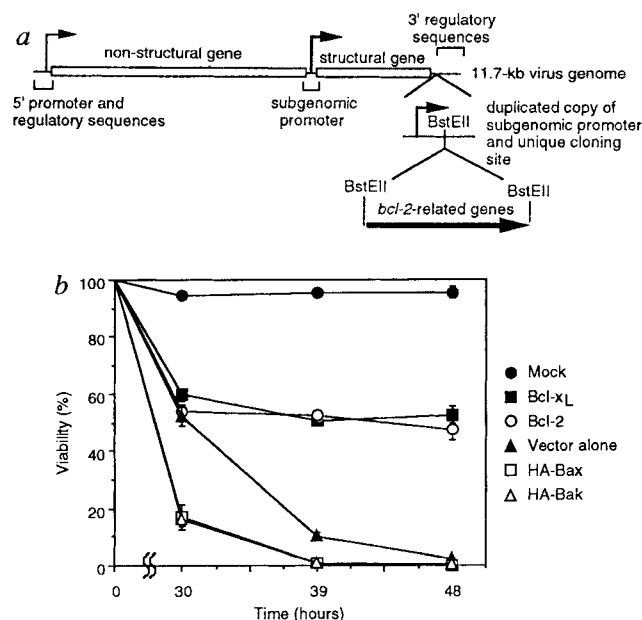
**THE Bcl-2-related protein, Bcl-x<sub>L</sub>, has been shown to block apoptosis induced by a variety of stimuli<sup>1-5</sup> and to be a stronger protector against apoptosis than Bcl-2 under certain circumstances<sup>2,5</sup>. Using site-specific mutagenesis, we show here that the amino-acid residues critical for protection of cells by Bcl-x<sub>L</sub> against Sindbis virus-induced apoptosis are clustered within the Bcl-2-homology regions 1 and 2 (BH1 and BH2 regions). The residues necessary for Bcl-x<sub>L</sub> function are not identical to those required for Bcl-2 function<sup>6</sup>. Although it has been suggested that heterodimerization between Bcl-x<sub>L</sub> and Bax is essential for the anti-death activity of Bcl-x<sub>L</sub> (refs 7, 8), our results suggest that the interaction with Bax is not required for Bcl-x<sub>L</sub> to exert its death-repressing activity. Specific mutations that disrupt the ability of Bcl-x<sub>L</sub> to interact with Bax or Bak still preserve 70–80% of the anti-death activity of wild-type Bcl-x<sub>L</sub>.**

Sindbis virus, an alphavirus, can induce apoptosis<sup>9-11</sup> which is inhibited by agents that block apoptosis induced by other stimuli<sup>9,10,12</sup>. Using the Sindbis virus genome as a vector to express death-regulatory genes, we devised a system to test rapidly the ability of candidate genes to suppress or accelerate virus-induced apoptosis (Fig. 1a). Bcl-2 and Bcl-x<sub>L</sub> expressed from the recombinant viruses protected BHK cells from apoptosis, as indicated by ~50% cell viability compared to 2% with virus vector alone at 48 hours post-infection (Fig. 1b). This is consistent with our earlier finding that a cell line stably transfected with *bcl-2* was protected from Sindbis virus-induced cell death<sup>9,10</sup>. Other members of the

Bcl-2 family, including Bax and Bak, have been shown to enhance cell death<sup>13-16</sup> and antagonize the protective effects of Bcl-2 in a dominant fashion<sup>13,15</sup>. Similarly, Bax and Bak expressed from recombinant viruses accelerated virus-induced apoptosis, as indicated by cell viabilities of 16% compared to 52% with the virus vector alone at 30 hours post-infection (Fig. 1b). Control recombination viruses with *bcl-2*-related genes inserted in the antisense orientation or with stop codons inserted into the anterior half of *bcl-2*-related genes were as competent as virus vector alone in triggering apoptosis (Fig. 1b and data not shown).

Although both Bcl-x<sub>L</sub> and Bcl-2 inhibit apoptosis triggered by a variety of stimuli in gene-transfer studies<sup>17-20</sup>, only Bcl-x<sub>L</sub> is required for embryonic development<sup>17-20</sup>. To investigate further the anti-death function of Bcl-x<sub>L</sub>, we engineered extensive site-specific mutagenesis in both BH1 and BH2 domains of human Bcl-x<sub>L</sub> (Fig. 2). Mutated *bcl-x<sub>L</sub>* genes were inserted into the Sindbis virus vector and tested for their ability to protect cells from Sindbis virus-induced apoptosis. Each recombinant virus clone expressed comparable levels of Bcl-x<sub>L</sub> proteins (Fig. 3d). Within the BH1 domain, amino-acid substitutions of the conserved residues Val-Asn-Trp or Gly-Arg-Ile (mutants 7 and 8) completely abolished the death-repressor activity. Replacing two of three phenylalanine residues with valine (mutants 9 and 10) partially impaired its protective effect. Single or double amino-acid substitutions between residues 129 and 134 (mutants 1–6) had minimal effects on the death-repressor activity, which differs from previous results with Bcl-2 (ref. 6). To verify that this difference is not due to apoptosis triggered by two different stimuli, amino acids were substituted in Bcl-2 at Phe 138 and at Asp 140 (corresponding to Bcl-x<sub>L</sub> mutant 1). This mutation totally abolished the anti-apoptotic effect of Bcl-2 in our virus vector system (unpublished results), consistent with the previous report<sup>6</sup>. Within the BH2 domain, amino-acid substitution of Trp 188 and Asp 189 (mutant 16) reduced the death-repressor activity of Bcl-x<sub>L</sub> by about half. Surprisingly, the Trp residue at position 181, which is highly conserved among all family members and critical for the death-repressor activity of Bcl-2 (ref. 6), can be replaced by glycine without affecting its activity (mutants 11–13). Substitution of Glu 193 (mutant 17) had a minimal effect on the death-repressor activity, consistent with the observation that Bcl-x<sub>L</sub> lacking residue 193 as a result of alternative splicing<sup>21</sup> can still protect neurons from death induced by nerve growth factor deprivation<sup>3</sup>. In contrast, the same substitution of Glu 200 in Bcl-2 abrogated half of the anti-apoptotic activity of Bcl-2, consistent with the observation that Bcl-2β is not as protective as full-length Bcl-2 (refs. 22–24).

**FIG. 1** Modulation of Sindbis virus-induced apoptosis by Bcl-2 related proteins. **a**, Diagram of double subgenomic Sindbis virus vector. **b**, Viability of BHK cells infected with the indicated recombinant Sindbis viruses at a multiplicity of infection (MOI) of 5 plaque-forming units (PFU) per cell. Data are presented as mean ± s.d. of 3 independent experiments. **METHODS.** The Sindbis virus vector was generated from plasmids TE12 (nucleotides 1–10,381; 11,382–13,637) and TZJSINC (nucleotides 10,382–11,485)<sup>26,27</sup> with *Bst*III restriction site inserted at the junction between nucleotide 11,485 of TZJSINC and nucleotide 11,382 of TE12. The resulting construct contains a cDNA of the Sindbis virus genome with a duplicated copy of the viral subgenomic promoter and a unique *Bst*III cloning site inserted between the viral coding sequences and the 3' untranslated/regulatory region of the Sindbis virus genome. Human *bcl-2*, human *bcl-x<sub>L</sub>*, human *bax* with an influenza virus haemagglutinin (HA) epitope, and human *bak* with the HA epitope were each inserted into the Sindbis virus vector. Following linearization, infectious RNA was transcribed *in vitro* from an SP6 promoter located immediately 5' of the viral genome sequence. Stocks of recombinant viruses were generated by transfecting the RNA into BHK cells and collecting the supernatant at 24 h post transfection. Cell viability was determined by trypan-blue exclusion.





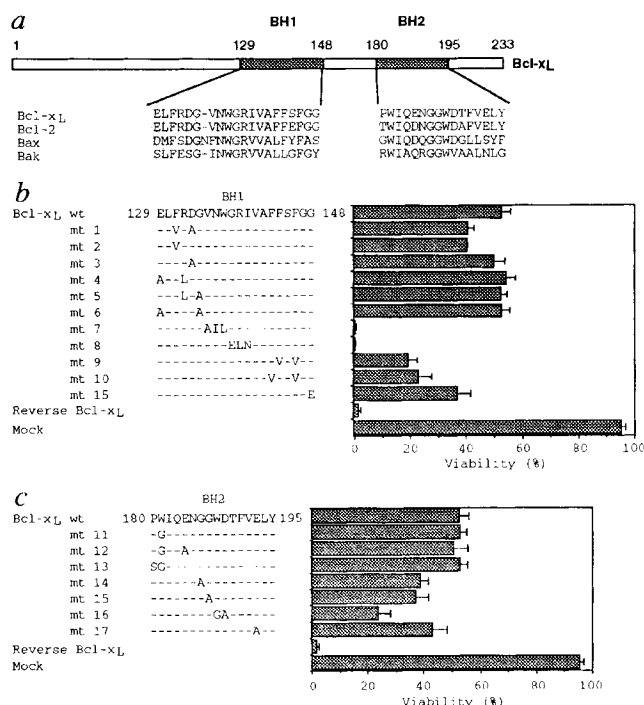
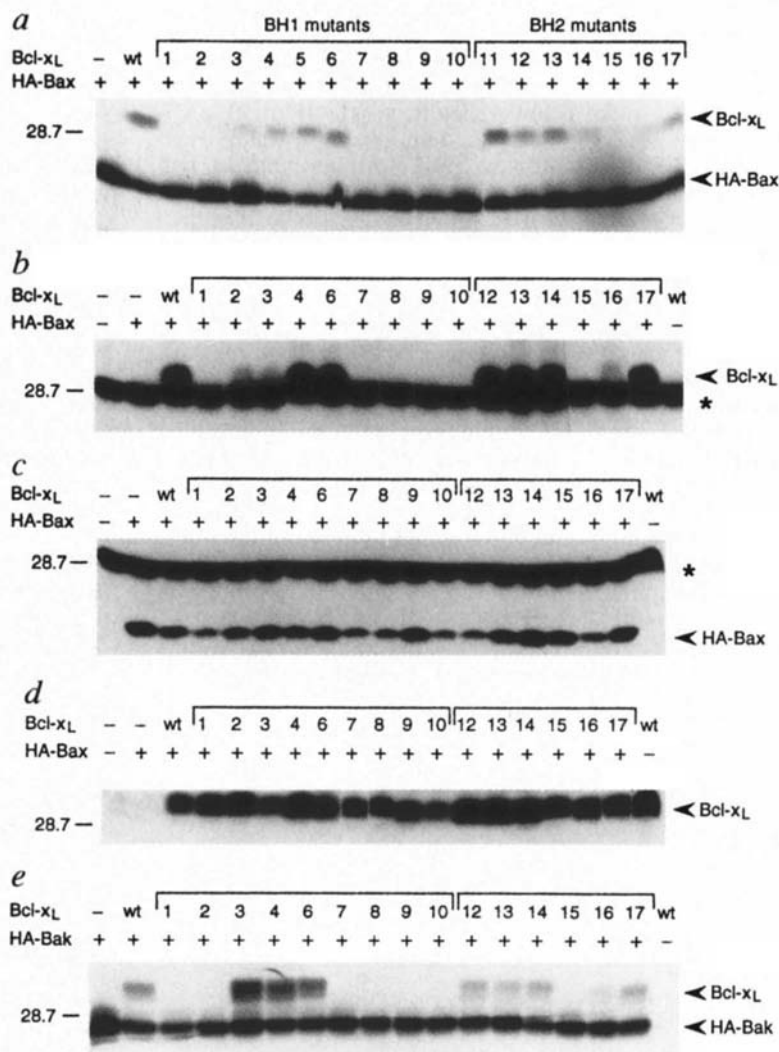


FIG. 2 Effect of BH1 and BH2 mutations on the death-repressor function of Bcl-xL. **a**, Alignment of amino acids from the conserved BH1 and BH2 domains of Bcl-2 family members<sup>6,16</sup>. **b** and **c**, Amino-acid substitutions in BH1 and BH2 domains of Bcl-xL; dashes indicate no changes.

METHODS. Oligonucleotide-directed mutagenesis was accomplished by two-step polymerase chain reaction (PCR). Mutants were cloned into pSG5 (Stratagene) and mutations were confirmed by dideoxy sequencing. DNA fragments containing the mutated *bcl-xL* genes were transferred from the pSG5 derivatives to the Sindbis virus vector. Recombinant viruses were generated as described in Fig. 1 legend and cell viability was determined at 48 h post-infection.

FIG. 3 Effect of Bcl-xL mutations on heterodimerization with Bax and Bak. **a**, Coprecipitation of *in vitro* translated Bcl-xL and HA-Bax with 12CA5 anti-HA antibody (Berkeley). **b**, Coprecipitation of Bcl-xL and HA-Bax in cell lysates with anti-HA antibody. BHK cells were co-infected with recombinant Sindbis viruses expressing HA-Bax and Bcl-xL or Bcl-xL mutants. Immunoprecipitates were immunoblotted with anti-Bcl-x antibody. **c**, A duplicate of gel shown in **b** immunoblotted with anti-HA antibody. The asterisk indicates the light chain of 12CA5 monoclonal anti-HA antibody used for immunoprecipitation. **d**, Bcl-xL expression in Sindbis virus vector-infected cells before immunoprecipitation. **e**, Coprecipitation of *in vitro* translated Bcl-xL and HA-Bak with anti-HA antibody. Migration position of  $M_r$  standard (28.7K) is shown.

METHODS. Bcl-xL or Bcl-xL mutants plus HA-Bax or HA-Bak were cotranslated *in vitro* with <sup>35</sup>S-methionine. Reticulocyte lysates (10  $\mu$ l) containing equivalent amounts of Bcl-xL and HA-Bax or HA-Bak (not shown), plus 290  $\mu$ l 0.2% NP-40 lysis buffer<sup>13</sup> were immunoprecipitated with 12CA5 anti-HA antibody and analysed by SDS-PAGE. 10<sup>6</sup> BHK cells were infected at a MOI of 10 PFU per cell for each virus and lysed with 0.2% NP-40 lysis buffer 10–12 h after infection. Cell lysates were immunoprecipitated with 12CA5, divided into equal portions, loaded onto duplicate 15% SDS-polyacrylamide gels and immunoblotted with 2A1 anti-human Bcl-x monoclonal antibody<sup>28</sup> or with 12CA5 using the ECL system (Amersham).



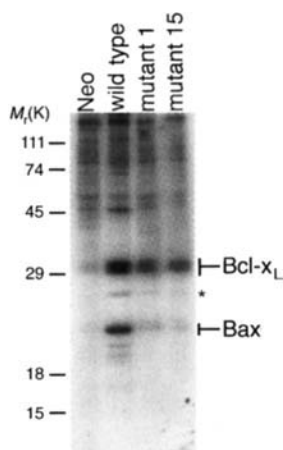


FIG. 4 Lack of interaction between Bcl- $x_L$  mutants and Bax/Bax-like proteins. The indicated stably transfected FL5.12 cells were prepared as described<sup>1</sup> and  $2 \times 10^7$  cells were labelled with 1 mCi of Trans  $^{35}\text{S}$ -label (ICN) for 8 h. Cell lysates were incubated for 1 h with mouse anti-human Bcl-x monoclonal antibody 7B2<sup>29</sup> (or IgG3 isotype control; not shown) and immunoprecipitates were analysed on 15% SDS-polyacrylamide gels treated with Amplify (Amersham) before autoradiography. The data are representative of 3 independent experiments. The asterisk marks the position of Bcl-2 that was coprecipitated with wild-type and mutant Bcl- $x_L$ .

To explore the role of heterodimerization between Bax and Bcl- $x_L$  in regulating cell death, we examined each Bcl- $x_L$  mutant for its association with haemagglutinin (HA)-Bax. Similar results were observed with proteins translated *in vitro* (Fig. 3a) as with protein complexes isolated from dually infected cell lysates (Fig. 3b). For about half of the Bcl- $x_L$  mutants tested, there was a correlation between the ability to heterodimerize with Bax and the ability to inhibit cell death. However, results from four mutants indicate that heterodimerization with Bax is not required for Bcl- $x_L$  to block apoptosis. Amino-acid substitution of either Phe 131 and Asp 133 (mutant 1) or Gly 148 and Gly 187 (mutant 15) completely abolished the interaction between Bcl- $x_L$  and Bax in coimmunoprecipitation assays (Fig. 3a, b), but still preserved 70–80% of the wild-type anti-apoptotic activity (Fig. 2b). Similarly, mutants 9 and 10, which could not bind Bax, still retained about half the activity of wild-type Bcl- $x_L$ . Furthermore, two additional mutants (mutants 2 and 3) were almost as protective as wild-type Bcl- $x_L$ , in spite of impaired binding to Bax. The total precipitated HA-Bax protein from cell lysates varied slightly but did not correlate with the observed differences in heterodimerization (Fig. 3c and data not shown). Another death-accelerating Bcl-2 homologue, Bak, which can heterodimerize with Bcl- $x_L$  (ref. 14), was also tested for its interaction with Bcl- $x_L$  mutants. The amino-acid residues of Bcl- $x_L$  involved in Bcl- $x_L$ /Bak heterodimerization are the same as those involved in Bcl- $x_L$ /Bax heterodimerization, except that mutation of Asp 133 impaired heterodimerization with Bax but not with Bak (Fig. 3e). FL5.12 cells stably transfected with mutants 1 and 15 were also protected from cell death following IL-3 withdrawal (48-h cell viabilities of 60.8 and 69.7%, respectively) compared to *neo* controls (7.8% viability) and wild-type Bcl- $x_L$  (87.1% viability). The reduced level of expression of mutant protein compared to wild-type Bcl- $x_L$  in FL5.12 cells may explain their slightly reduced protection (Fig. 4). An experiment similar to that shown in Fig. 3b also failed to detect heterodimerization between Bax and Bcl- $x_L$  mutants 1 and 15 in

FL5.12 cells (data not shown). To verify that mutants 1 and 15 do not bind other as-yet unidentified Bax-like proteins present in the cell,  $^{35}\text{S}$ -labelled cell lysates from FL5.12 cells stably transfected with mutant 1 or 15 were immunoprecipitated with anti-Bcl-x antibodies. Although wild-type Bcl- $x_L$  coprecipitated with endogenous Bax protein, both mutants failed to coprecipitate with Bax or Bax-like proteins (Fig. 4). The amount of Bax in the mutant lanes corresponds to that coprecipitated with endogenous Bcl- $x_L$  in control *neo* cells.

We have presented evidence that BH1 and BH2 domains are both critical for the death-repressor function of Bcl- $x_L$ . Substitution of conserved amino acids within these domains suggests that the residues required for Bcl- $x_L$  function may be different in Bcl-2 (ref. 6), which could contribute to the specificity of these two proteins. More important, the ability to interact with Bax or Bak is neither required for nor closely correlates with the ability of Bcl- $x_L$  to exert its protective activity against an apoptotic stimulus.

It has been debated whether the death-suppressing Bcl-2 homologues or the death-accelerating members are the key effectors in modulating the cell death pathway. If the death-suppressing Bcl-2 homologues function as anti-death effectors, then the death-accelerating homologues may block the protective effects of Bcl-2 and Bcl- $x_L$  through sequestration or competition for downstream targets. Alternatively, if the death-accelerating Bcl-2 homologues generate death signals, then the death-suppressing homologues may serve as dominant inhibitors. Although the findings with Bax-deficient mice suggest that these models are not mutually exclusive<sup>25</sup>, our data support the former model. It has been reported that Bcl-2 must bind to Bax to exert its anti-death function. Therefore, the working mechanism for Bcl- $x_L$  may be different from Bcl-2. However, it is equally likely that the domain of Bcl-2 required to heterodimerize with Bax totally overlaps with the binding site for a downstream target. Thus, identification of new proteins that bind to Bcl- $x_L$  mutants 1 and 15 should provide insight into the mechanism by which Bcl- $x_L$  blocks apoptosis. □

Received 20 September; accepted 5 December 1995.

- Boise, L. H. et al. *Cell* **74**, 597–608 (1993).
- Gottschalk, A. R., Boise, L. H., Thompson, C. B. & Quintans, J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7350–7354 (1994).
- González-García, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 4304–4308 (1995).
- Dole, M. G. et al. *Can. Res.* **55**, 2576–2582 (1995).
- Shimizu, S. et al. *Nature* **374**, 811–813 (1995).
- Yin, X.-M., Oltvai, Z. N. & Korsmeyer, S. J. *Nature* **369**, 321–323 (1994).
- Yang, E. et al. *Cell* **80**, 285–291 (1995).
- Sedlak, T. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 7834–7838 (1995).
- Levine, B. et al. *Nature* **361**, 739–742 (1993).
- Ubol, S., Tucker, P. C., Griffin, E. E. & Hardwick, J. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5202–5206 (1994).
- Lewis, J., Wesselring, S. L., Griffin, D. E. & Hardwick, J. M. *J. Virol.* **70**, 1828–1835.
- Lin, K.-I. et al. *J. Cell Biol.* **131**, 1149–1161 (1995).
- Oltvai, Z. N., Millman, C. L. & Korsmeyer, S. J. *Cell* **74**, 609–619 (1993).
- Farrow, S. N. et al. *Nature* **374**, 731–733 (1995).
- Chittenden, T. et al. *Nature* **374**, 733–736 (1995).
- Kiefer, M. C. et al. *Nature* **374**, 736–739 (1995).

- Nakayama, K.-I. et al. *Science* **261**, 1584–1588 (1993).
- Veis, D. J., Sorenson, C. M., Shutter, S. R. & Korsmeyer, S. J. *Cell* **75**, 229–240 (1993).
- Kamada, S. et al. *Can. Res.* **55**, 354–359 (1995).
- Motoyama, N. et al. *Science* **267**, 1506–1510 (1995).
- González-García, M. et al. *Development* **120**, 3033–3042 (1994).
- Tanaka, S., Saito, K. & Reed, J. C. *J. Biol. Chem.* **268**, 10920–10926 (1993).
- Hockenbery, D. M., Oltvai, Z. N., Yin, X.-M., Millman, C. L. & Korsmeyer, S. J. *Cell* **75**, 241–251 (1993).
- Borner, C. et al. *J. Cell Biol.* **126**, 1059–1068 (1994).
- Knudson, C. M., Tung, K. S. K., Tourtellotte, W. G. & Brown, G. A. J. & Korsmeyer, S. J. *Science* **270**, 96–99 (1995).
- Lustig, S. et al. *J. Virol.* **62**, 2329–2336 (1988).
- Hertz, J. M. & Huang, H. V. *J. Virol.* **66**, 857–864 (1992).
- Boise, L. H., Minn, A. J., Noel, P. J., June, C. H. & Thompson, C. B. *Immunity* **3**, 87–98 (1995).
- Gottschalk, A. R. et al. *Cell Death and Differentiation* (in the press).

ACKNOWLEDGEMENTS. We thank M. Veliou for technical assistance. This work was supported by grants from the Multiple Sclerosis Society and the NIH (J.M.H.). L.H.B. is a grantee of the Leukemia Society of America.

# Role of transactivation of the EGF receptor in signalling by G-protein-coupled receptors

Henrik Daub, F. Ulrich Weiss, Christian Wallasch & Axel Ullrich\*

Department of Molecular Biology, Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, 82152 Martinsried, Germany

**TRANSDUCTION of a mitogenic signal from the cell membrane to the nucleus involves the adapter proteins SHC and Grb2, which mediate activation of the Ras/mitogen-activated protein (MAP) kinase pathway<sup>1-5</sup>. In contrast to receptor tyrosine kinases (RTKs), the signalling steps leading to Ras/MAP kinase activation by G-protein-coupled receptors (GPCRs) are still poorly characterized but appear to include  $\beta\gamma$  subunits of heterotrimeric G-proteins and as-yet unidentified tyrosine kinases<sup>6-8</sup>. We report here that the epidermal growth factor receptor (EGFR) and the *neu* oncoprotein become rapidly tyrosine-phosphorylated upon stimulation of Rat-1 cells with the GPCR agonists endothelin-1, lysophosphatic acid and thrombin, suggesting that there is an intracellular mechanism for transactivation. Specific inhibition of EGFR function by either the selective tyrphostin AG1478 or a dominant-negative EGFR mutant suppressed MAP kinase activation and strongly inhibited induction of *fos* gene expression and DNA synthesis. Our results demonstrate a role for RTKs as downstream mediators in GPCR mitogenic signalling and suggest a ligand-independent mechanism of RTK activation through intracellular signal crosstalk.**

Analysis of endogenous GPCR signalling events in Rat-1 fibroblasts led to the identification of a major tyrosine-phosphorylated glycoprotein of relative molecular mass 170K in response to stimulation with endothelin-1 (ET-1), lysophosphatidic acid (LPA) and thrombin, factors that induce mitogenic stimuli in these cells<sup>9-11</sup>. We identified this protein as the epidermal growth factor receptor (Fig. 1a). Endothelin-induced tyrosine phosphorylation of EGFR was inhibited by BQ123, a selective antagonist for the ET<sub>A</sub> receptor, which represents the predominant subtype in Rat-1 cells<sup>12</sup>. Thrombin-receptor peptide (TRP) was able to mimic the thrombin response, confirming specific activation of the thrombin GPCR<sup>13</sup>. In comparison with GPCR-mediated stimulation, EGF-induced receptor autophosphorylation was more pronounced, whereas the vehicle BSA, included as an LPA control, and platelet-derived growth factor (PDGF), showed no effect (Fig. 1a). Time course experiments revealed rapid transient tyrosine phosphorylation of the EGFR, reaching a maximum within 2 min after stimulation, as shown representatively for ET-1 (Fig. 1b, and results not shown). Owing to the rapid kinetics, the involvement of EGFR ligands in this receptor-mediated signal activation process must be excluded, pointing instead to the release of a negative restraint mechanism on RTK activity that probably involves a protein-tyrosine phosphatase.

Rat-1 cells were then infected with a retrovirus designed to express human EGFR (HER1)<sup>14</sup>. Immunoprecipitation with monoclonal antibody 108.1, which selectively recognizes the human but not rat EGFR, demonstrated ET-1-stimulated tyrosine phosphorylation of ectopically expressed HER1 (Fig. 1c). Moreover, GPCR-mediated RTK activation was not exclusive for the EGFR, for we also detected increased tyrosine phosphorylation of endogenous rat p185<sup>neu</sup> (ref. 15) upon stimulation with ET-1, LPA or thrombin (Fig. 1d). As the induced phosphotyrosine signal was relatively weaker than the EGFR response, we cannot exclude

FIG. 1 GPCR stimulation induces tyrosine phosphorylation of EGFR and p185<sup>neu</sup> in Rat-1 fibroblasts. **a**, Quiescent cells were stimulated for 5 min with ligands for GPCRs (100 nM ET-1, 25  $\mu$ M LPA, 2 U per ml thrombin, 50  $\mu$ M TRP) or RTKs (10 ng ml<sup>-1</sup> EGF, 30 ng ml<sup>-1</sup> PDGF-B/B) and lysed. Immunoprecipitations (IP) were done with anti-EGFR ( $\alpha$ EGFR) antibody (from UBI). BQ123 (2  $\mu$ M) preincubation was for 10 min. **b**, Induction of EGFR tyrosine phosphorylation was analysed upon stimulation of Rat-1 cells for 2, 5, 10, 20 and 30 min with 100 nM ET-1. **c**, Cell lines Rat-1-pLXSN and Rat-1-HER1 were stimulated with 100 nM ET-1. Ectopically expressed HER1 was precipitated with the human EGFR-specific mAb 108.1 (ref. 14). **d**, Induction of p185<sup>neu</sup> tyrosine phosphorylation was analysed after 5 min stimulation of Rat-1 cells with GPCR agonists or RTK ligands. Lysates were precipitated with a polyclonal rabbit antiserum against the C-terminal dodecapeptide of the receptor. Molecular size is indicated on the right.

**METHODS.** Rat-1 cells were grown to 90% confluency in DMEM medium containing 5% fetal bovine serum, serum-starved for 24 h and treated for the indicated times with appropriate ligands. Cells were lysed and immunoprecipitated as described<sup>14</sup>. Precipitates were analysed by SDS-PAGE (7.5% gels) and transferred to nitrocellulose. Immunoblots were probed with 5E2 antiphosphotyrosine antibodies<sup>14</sup> and developed with an enhanced chemiluminescence (ECL Amersham) detection system. Before reprobing, filters were stripped according to manufacturer's protocol. For generation of Rat-1-pLXSN, Rat-1-HER1 and Rat-1-HERCD533 cell lines, subconfluent cells (10<sup>5</sup> cells per 6-cm dish) were repeatedly incubated with pLXSN, pLXSN-HER1, or pNTK-HERCD533 virus supernatants of GP + E 86 cells (> 1  $\times$  10<sup>6</sup> G418 CFU ml<sup>-1</sup>; MOI of 10) for 4 to 12 h.

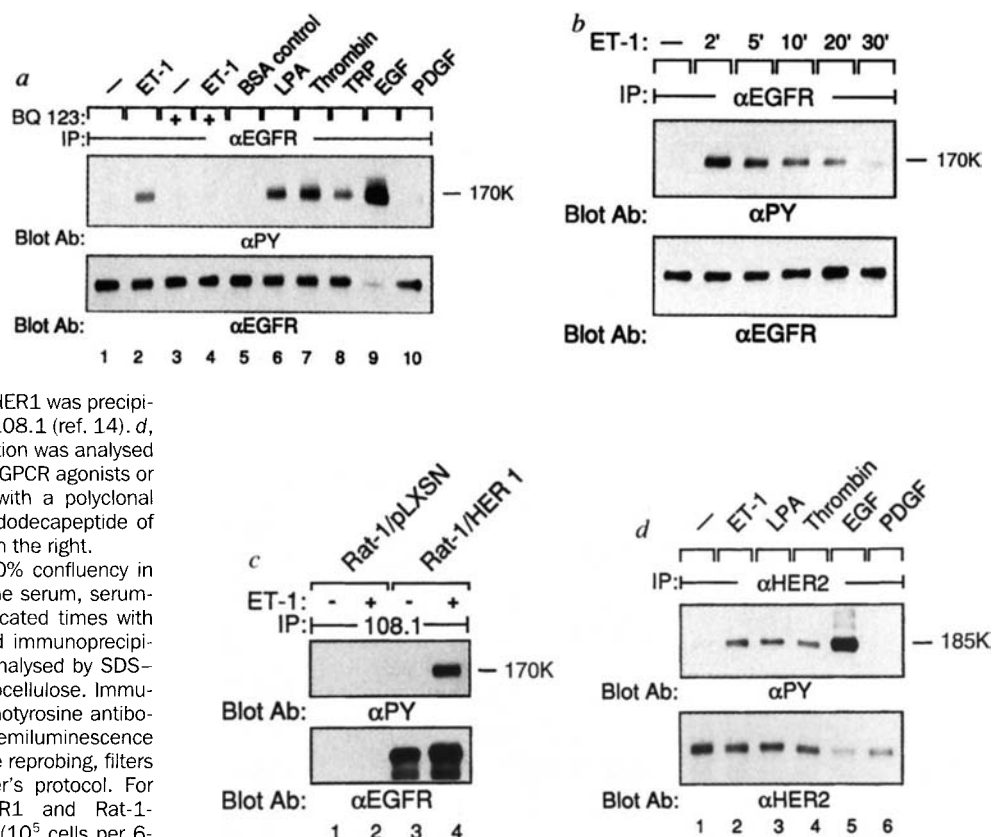


FIG. 2 Specific inhibition of GPCR agonist-induced EGFR tyrosine phosphorylation. **a**, Expression of the dominant-negative EGFR mutant HERCD533 (ref. 14) in Rat-1-HERCD533 cells was confirmed by [ $^{35}$ S]methionine metabolic labelling, immunoprecipitation with mAb 108.1, and SDS-PAGE followed by autoradiography. **b**, Rat-1-pLXSN and Rat-1-HERCD533 cells were stimulated for 5 min with indicated ligands (100 nM ET-1, 25  $\mu$ M LPA, 2 U ml $^{-1}$  thrombin, 2 and 10 ng ml $^{-1}$  EGF). Upon lysis, precipitated EGFR was immunoblotted with 5E2 antibodies and reprobed for EGFR. **c**, Quiescent Rat-1 cells were stimulated with indicated ligands (100 nM ET-1, 2 U ml $^{-1}$  thrombin, 10 ng ml $^{-1}$  EGF) following preincubation of cells with or without EGRF-specific tyrosine phosphatase AG1478 (250 nM, 15 min). Lysates were processed as for **b**. Molecular size is indicated on the right. **d**, Top, anti-paxillin immunoprecipitates from the same cell lysates as in **c**, analysed for tyrosine phosphorylation and reprobed with monoclonal anti-paxillin antibody (from Signal Transduction). Bottom, tyrosine phosphorylation of FAK was detected in whole-cell lysates. Equal expression was confirmed by reprobating with an anti-FAK-specific antibody (Santa Cruz Biotech.).

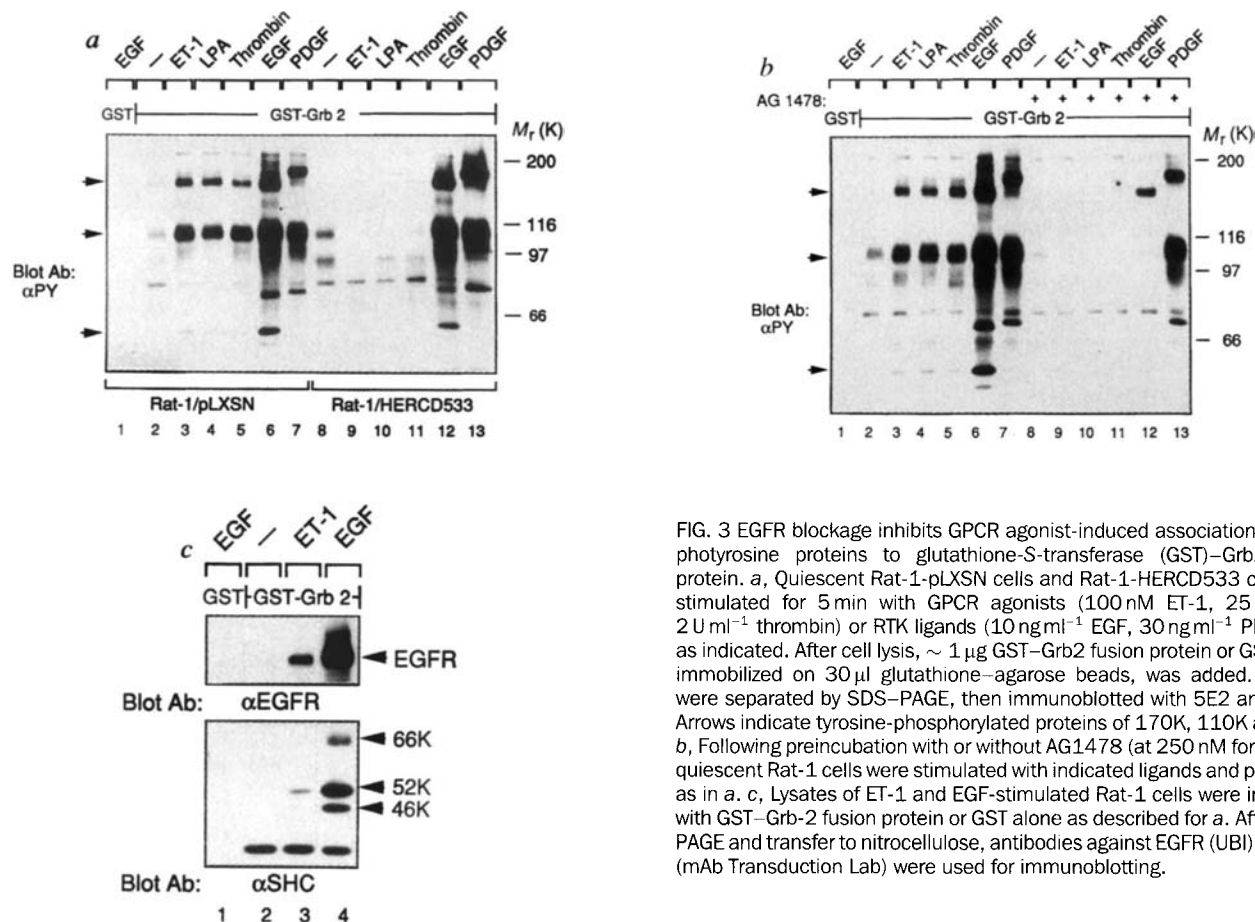
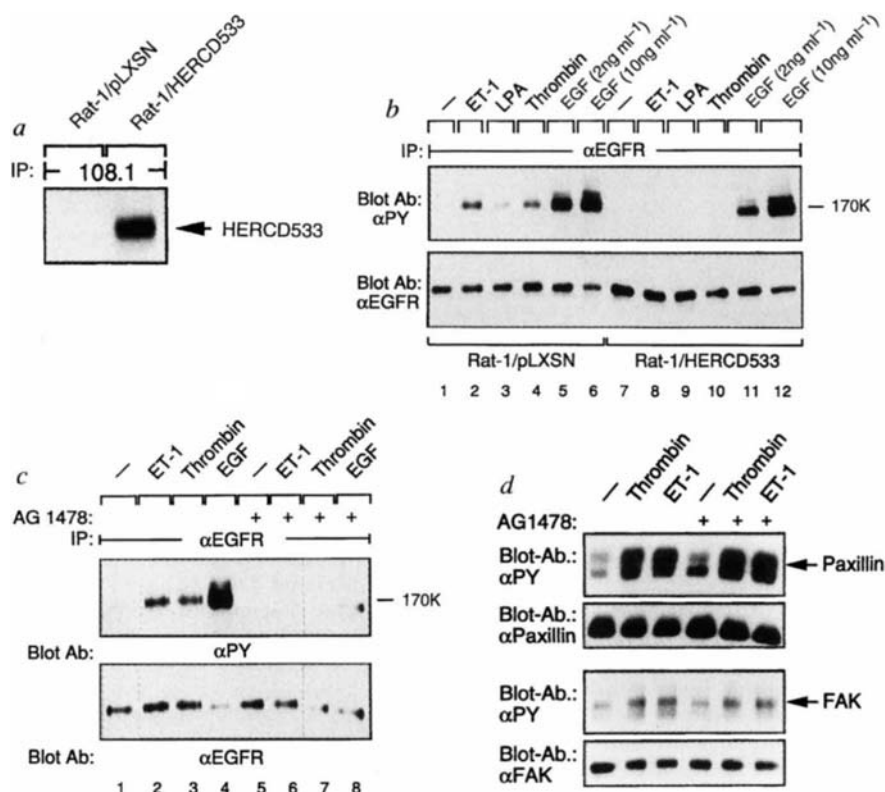


FIG. 3 EGFR blockage inhibits GPCR agonist-induced association of phosphotyrosine proteins to glutathione-S-transferase (GST)-Grb2 fusion protein. **a**, Quiescent Rat-1-pLXSN cells and Rat-1-HERCD533 cells were stimulated for 5 min with GPCR agonists (100 nM ET-1, 25  $\mu$ M LPA, 2 U ml $^{-1}$  thrombin) or RTK ligands (10 ng ml $^{-1}$  EGF, 30 ng ml $^{-1}$  PDGF-B/B) as indicated. After cell lysis,  $\sim$ 1  $\mu$ g GST-Grb2 fusion protein or GST alone, immobilized on 30  $\mu$ l glutathione-agarose beads, was added. Proteins were separated by SDS-PAGE, then immunoblotted with 5E2 antibodies. Arrows indicate tyrosine-phosphorylated proteins of 170K, 110K and 52K. **b**, Following preincubation with or without AG1478 (at 250 nM for 10 min), quiescent Rat-1 cells were stimulated with indicated ligands and processed as in **a**. **c**, Lysates of ET-1 and EGF-stimulated Rat-1 cells were incubated with GST-Grb2 fusion protein or GST alone as described for **a**. After SDS-PAGE and transfer to nitrocellulose, antibodies against EGFR (UBI) and SHC (mAb Transduction Lab) were used for immunoblotting.



that p185<sup>neu</sup> phosphorylation may result from EGFR transphosphorylation.

The role of EGFR in GPCR signalling was further analysed by using two independent approaches for specific inhibition of the EGFR signal. First the dominant-negative EGFR mutant HERCD533, which lacks the cytoplasmic domain, was stably transfected in Rat-1 cells (Fig. 2a). This mutant inhibits EGFR downstream signalling by formation of signalling-defective heterodimers with the wild-type receptor<sup>14,16</sup>. At low ligand concentrations, receptor activation was only attenuated compared to control infected cells, whereas tyrosine phosphorylation of the EGFR following GPCR stimulation was completely blocked (Fig. 2b). Second, pretreatment of cells with tyrphostin AG1478, a selective EGFR antagonist<sup>17</sup>, inhibited receptor tyrosine phosphorylation following stimulation with EGF or GPCR ligands (Fig. 2c). In both experiments, GPCR-induced tyrosine phosphorylation of focal adhesion kinase and paxillin<sup>18</sup> was unaffected (Fig. 2d, and results not shown), demonstrating selective inhibition of EGFR activation which does not interfere with functional coupling to other GPCR downstream signalling targets.

Association of tyrosine-phosphorylated proteins with Grb2 is known to represent a critical step in the EGFR signalling cascade. Following treatment of Rat-1 cells with ET-1, LPA or thrombin, tyrosine-phosphorylated proteins of 170K, 110K and 52K were found to associate with a glutathione-S-transferase (GST) Grb2 fusion protein. In contrast, in cells where EGFR activity was blocked either by expression of the HERCD533 mutant (Fig. 3a) or by AG1478 pretreatment (Fig. 3b), these proteins did not associate. The 170K and 52K phosphoproteins were identified as EGFR and SHC (Fig. 3c), but the 110K protein was not identified. SHC and the 110K protein are apparently downstream targets of 'transactivated' EGFR. Several GPCR agonists have been shown to induce SHC tyrosine phosphorylation and association with Grb2 (refs 19–21). Our identification of the EGFR as an essential mediator of SHC tyrosine phosphorylation following GPCR stimulation strongly suggests that analogous signalling mechanisms through transactivation of EGFR, p185<sup>neu</sup> or other RTKs may be used by GPCR ligands in diverse cell types.

As Grb2 association and phosphorylation events are essential in triggering the Ras/MAP kinase pathway, we investigated the

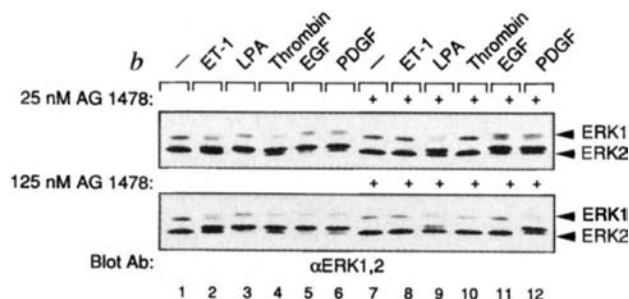
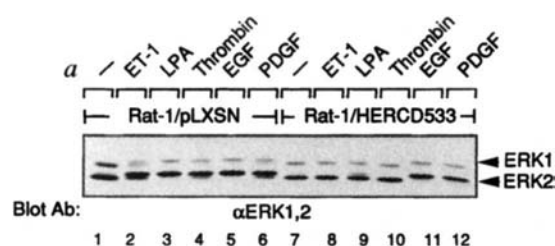


FIG. 4 EGFR inactivation inhibits

GPCR-mediated MAP kinase activation, c-fos induction and DNA synthesis.

a, b, MAP kinase gel shift assays.

a, Rat-1-HERCD533 and Rat-1-pLXSN cells were stimulated with the indicated ligands for 5 min.

b, Rat-1 cells were preincubated with the indicated concentrations of AG1478 (ref. 17) or dimethylsulphoxide (DMSO) for 5 min, followed by a further 5 min stimulation with the appropriate ligands.

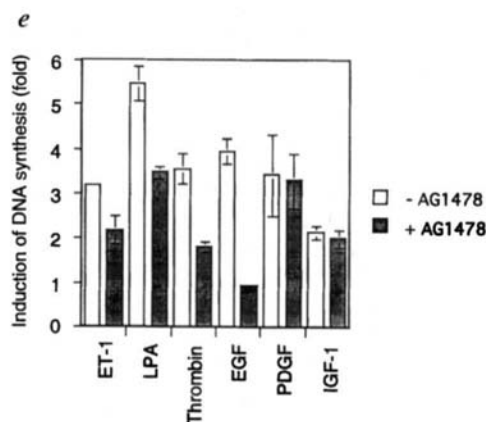
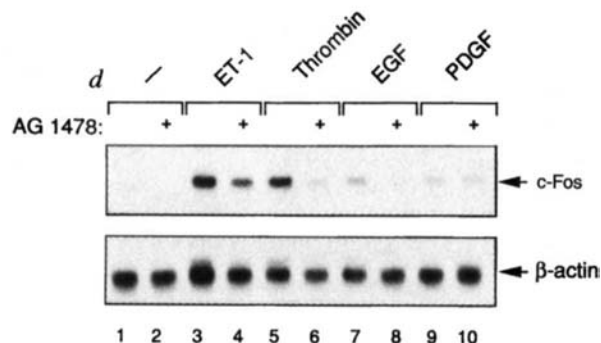
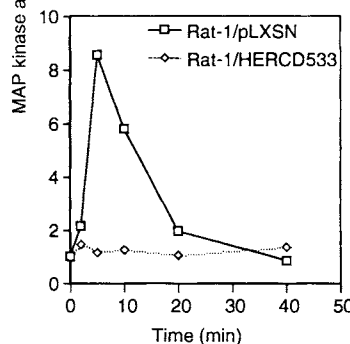
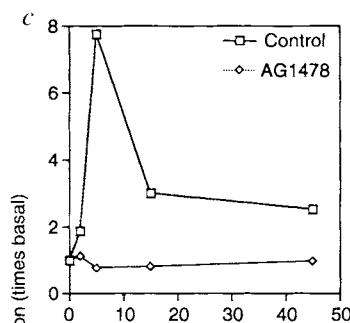
Cell lysates (20 µg) were analysed by immunoblotting with anti-ERK1,2 antibodies (Santa Cruz).

c, In-gel MAP kinase assay. Top, Control DMSO- or AG1478 (125 nM)-pretreated cells were stimulated with thrombin and MAP kinase activation at the indicated times was determined<sup>22</sup>.

Bottom, MAP kinase activation of Rat-1-pLXSN and Rat-1-HERCD533 cells was assayed at the indicated times after ET-1 stimulation.

d, Northern blot analysis of c-fos induction. Total RNA was extracted after 30 min treatment with the appropriate ligands in the presence or absence of tyrphostin AG1478 (250 nM, 5 min preincubation); 10 µg of each were separated on a 1.4% agarose formaldehyde gel and transferred to nitrocellulose for hybridization with a <sup>32</sup>P-labelled cDNA probe of c-fos full coding sequence. The membrane was rehybridized with a β-actin cDNA probe. Blots were exposed using an intensifying screen.

e, Incorporation of <sup>3</sup>H-thymidine into DNA. Rat-1 fibroblasts were seeded into 12-well plates (7 × 10<sup>4</sup> cells per cell). Upon serum deprivation for 24 h, cells were subjected to 5 min preincubation with



either 250 nM AG1478 (filled bars) or DMSO (open bars) before ligand treatment. After 18 h incubation, cells were pulse-labelled with <sup>3</sup>H-thymidine for 4 h, and thymidine incorporation measured by trichloroacetic acid precipitation and subsequent liquid-scintillation counting. Representative data of one of three experiments are shown; values are means ± s.d.

effects of EGFR inhibition on GPCR-induced activation of the MAP kinases ERK-1 and ERK-2. As determined by decreased electrophoretic mobility due to phosphorylation, ERK1/2 were activated following stimulation with ET-1, LPA and thrombin, as with control RTK ligands EGF and PDGF (Fig. 4a, b). In HERCD533 transfectants, however, ERK activation by ET-1 and thrombin was blocked and the LPA response was attenuated whereas activation by PDGF or EGF remained unaltered (Fig. 4a). In the latter case, the dominant-negative inhibitory effect of HERCD533 was overcome by stimulation with relatively high ligand concentrations ( $10 \text{ ng ml}^{-1}$ ) (Fig. 4a, lanes 11, 12), confirming previous observations<sup>14</sup>. Upon preincubation with AG1478, GPCR-ligand-triggered ERK activation was blocked at even lower concentrations than those necessary to block the EGF response (Fig. 4b). PDGF-induced ERK activation remained unaffected, confirming selective tyrphostin inhibition. In-gel measurement of MAP kinase activity revealed a fast transient induction upon ET-1 or thrombin stimulation, which could be completely blocked by AG1478 preincubation or HERCD533 expression (Fig. 4c). Based on these results, the EGFR must be considered to be an integral and essential element of the GPCR signalling cascade leading to MAP-kinase activation in Rat-1 fibroblasts.

We further investigated the contribution of EGFR activation to ET-1 and thrombin-induced *c-fos* expression. AG1478 pre-treatment significantly reduced *c-fos* induction upon ET-1 stimulation and more strongly upon thrombin stimulation, whereas there was no reduction of the PDGF response (Fig. 4d). Differential effects on *c-fos* gene expression may be due to synergistic signals involving, to different extents, Rho-dependent pathways<sup>22</sup>, *c-jun* N-terminal kinase<sup>23,24</sup> or other, so-far-unidentified, parallel systems.

For quantification of EGFR-transmitted mitogenic signalling, we measured GPCR-induced DNA synthesis in the presence or absence of the EGFR tyrphostin. AG1478 incubation completely blocked EGF-mediated induction of DNA synthesis, whereas ET-1, LPA and thrombin stimulation were reduced by 47, 45 and 69%, respectively (Fig. 4e); there was no significant effect on the mitogenic response to PDGF or IGF-1. We conclude that for the induction of a full proliferative response to ET-1, LPA or thrombin, functional EGFR signalling is required. The incomplete inhibition of GPCR-induced DNA synthesis by AG1478 may result from induction of secondary proliferative stimuli bypassing the blocked EGFR during the 22-hour assay or from other experimental factors.

Our findings indicate that mitogens such as ET-1, LPA and thrombin mediate cell proliferation through ligand-independent induction of tyrosine phosphorylation of the RTKs EGFR and p185<sup>neu</sup> in Rat-1 fibroblasts. It appears that additional tyrosine kinases contribute in a general or cell-type-specific way to GPCR-mediated mitogenic signalling, as indicated already by the involvement of the PDGF receptor in angiotensin II signalling in vascular smooth muscle cells<sup>25</sup>. The existence of synergistic parallel pathways that are activated by GPCR ligands is suggested by quantitative differences in the mitogenic signal of GPCR agonists upon EGFR inhibition. Differential recruitment of RTKs by different GPCRs, or even by the same GPCR in different cell types, opens a new dimension of biological signal diversity and cellular response definition through possible utilization of additional RTK pathways or regulatory systems involving signal-transducing factors such as phospholipase- $\gamma$ , protein tyrosine phosphatase 1D, phosphatidylinositol-3-OH kinase and others<sup>26-28</sup>.

The mechanism described here is reminiscent of that involving growth-factor receptors in cellular responses triggered by stress factors such as ultraviolet light or hydrogen peroxide. Like the effects of GPCR ligands, ultraviolet irradiation results in transient EGFR tyrosine phosphorylation and EGFR-mediated *c-fos* induction<sup>29</sup>. More details of this RTK transactivation mechanism are needed to clarify its role in the regulation of biological processes and the pathophysiology of diseases such as cancer. □

Received 27 September; accepted 12 December 1995.

1. Ullrich, A. & Schlessinger, J. *Cell* **61**, 203–212 (1990).
2. van der Geer, P. & Hunter, T. A. *Rev. Cell Biol.* **10**, 251–337 (1994).
3. Burgering, B. M. T. & Bos, J. L. *Trends biochem. Sci.* **20**, 18–22 (1995).
4. Egan, S. E. et al. *Nature* **363**, 45–51 (1993).
5. van Blesien, T. et al. *Nature* **376**, 781–784 (1995).
6. van Conven, E. J., Hordijk, P. L., Medema, R. H., Bos, J. L. & Moolenaar, W. H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1257–1261 (1993).
7. Faure, M., Voinyo-Yasenetskaya, T. A. & Bourne, H. J. *biol. Chem.* **269**, 7851–7854 (1994).
8. Crespo, P., Xu, N., Simonds, W. F. & Gutkind, J. A. *Nature* **369**, 418–420 (1994).
9. Muldoon, L. L., Pribnow, D., Rodland, K. D. & Magun, B. E. *Cell Reg.* **1**, 379–390 (1990).
10. van Conven, E. J., Groenink, A., Jalink, K., Eichholtz, T. & Moolenaar, W. H. *Cell* **59**, 45–54 (1989).
11. Gupta, S. K., Gallego, C., Johnson, G. L. & Heasley, L. E. *J. biol. Chem.* **267**, 7987–7990 (1992).
12. Ambar, I. & Sokolovsky, M. *Eur. J. Pharmac.* **245**, 31–41 (1993).
13. Vu, T.-K. H., Hung, D. T., Wheaton, V. I. & Coughlin, S. R. *Cell* **64**, 1057–1068 (1991).
14. Redemann, N., Holzmann, B., von Rüden, T., Wagner, E. F., Schlessinger, J. & Ullrich, A. *Molec. cell. Biol.* **12**, 491–498 (1992).
15. Bargmann, C. I., Hung, M.-C. & Weinberg, R. A. *Nature* **319**, 226–229 (1986).
16. Kashles, O., Yarden, Y., Fischer, R., Ullrich, A. & Schlessinger, J. *J. Molec. cell. Biol.* **11**, 1454–1463 (1991).
17. Levitzki, A. & Gazit, A. *Science* **267**, 1782–1788 (1995).
18. Zachary, I., Sinnett-Smith, J., Turner, C. E. & Rozengurt, E. *J. biol. Chem.* **268**, 22060–22065 (1993).
19. Ohmichi, M. et al. *J. biol. Chem.* **269**, 3783–3788 (1994).
20. Cazaubon, S. M. et al. *J. biol. Chem.* **269**, 24805–24809 (1994).
21. Lev, S. et al. *Nature* **276**, 737–745 (1995).
22. Hill, C. S., Wynne, J. & Treisman, R. *Cell* **81**, 1159–1170 (1995).
23. Cano, E., Hazzalin, C. A. & Mahadevan, L. C. *Molec. cell. Biol.* **14**, 7352–7362 (1994).
24. Coso, O. A. et al. *J. biol. Chem.* **270**, 5620–5624 (1995).
25. Linseman, D. A., Benjamin, C. W. & Jones, D. A. *J. biol. Chem.* **270**, 12563–12568 (1995).
26. Marrero, M. B., Paxton, W. G., Duff, J. L., Berk, B. C. & Bernstein, K. E. *J. biol. Chem.* **269**, 10935–10939 (1994).
27. Rivard, N., McKenzie, F. R., Brondello, J.-M. & Pouyssegur, J. *J. biol. Chem.* **270**, 11017–11024 (1995).
28. Kumagai, N., Morii, N., Fujisawa, K., Nemoto, Y. & Narumiya, A. *J. biol. Chem.* **268**, 24535–24538 (1993).
29. Sachsenmeier, C. et al. *Cell* **78**, 963–972 (1994).

ACKNOWLEDGEMENTS. H.D. and F.U.W. contributed equally to this work. We thank R. Lammers for the pGEX-Grb2 construct, P. Hirth and his colleagues for AG1478, G. Niederfellner for discussion, and J. Arch for preparing the manuscript. This work was supported by a grant from Sugen, Inc.

## A Grb2-associated docking protein in EGF- and insulin-receptor signalling

Marina Holgado-Madruga\*, David R. Emlet†, David K. Moscatello\*, Andrew K. Godwin‡ & Albert J. Wong\*†§

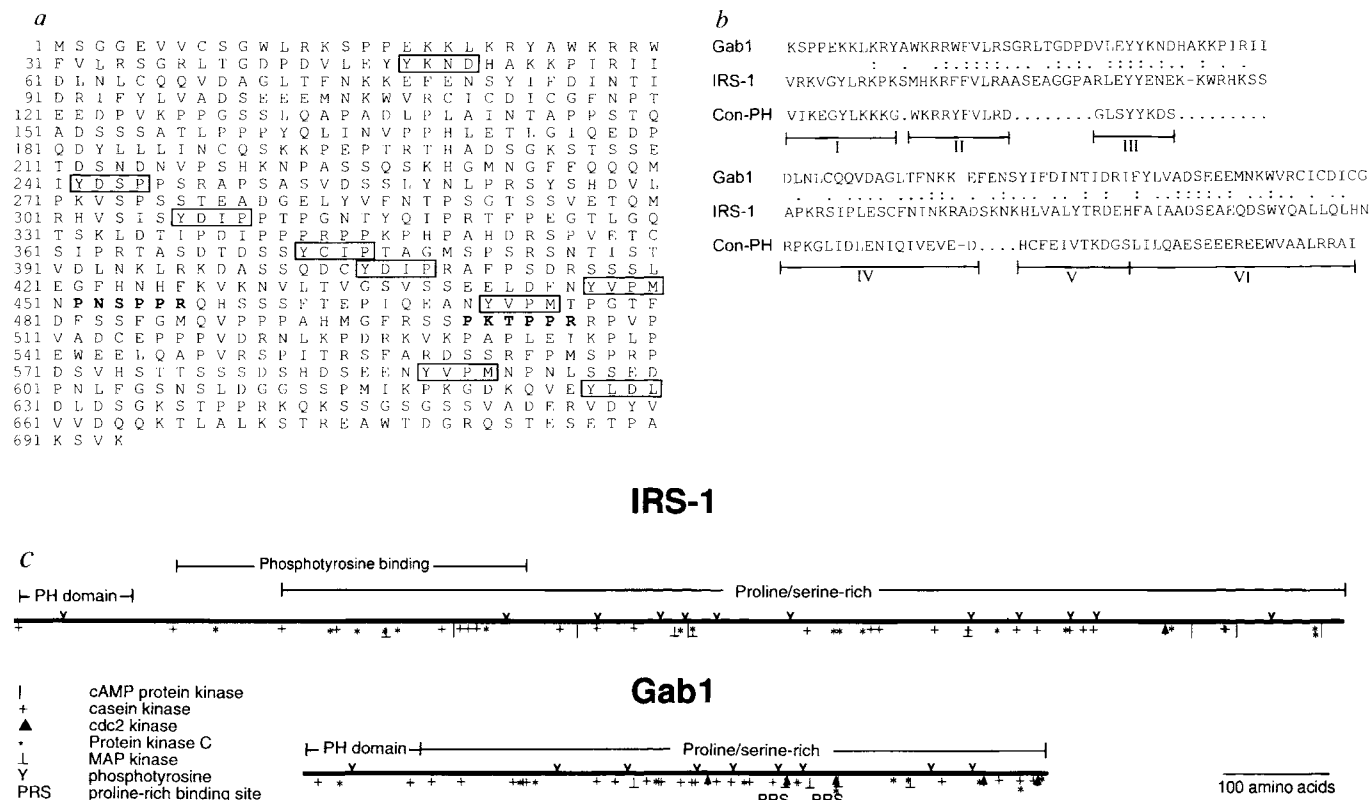
Departments of \*Microbiology & Immunology and †Pharmacology, Jefferson Cancer Institute, Philadelphia, Pennsylvania 19107, USA  
‡Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

The protein Grb2 plays a central role in signalling by receptor protein-tyrosine kinases<sup>1,2</sup>, where its SH2 domain binds to the receptor and its two SH3 domains link to effectors. One target effector is Sos, so Grb2 links receptor protein-tyrosine kinases with the Ras signalling pathway. The SH3 domains can also couple to other signalling proteins, including Vav<sup>3</sup>, c-Abl<sup>4</sup> and dynamin<sup>5</sup>. We have identified several bands in glial and medulloblastoma tumours that are recognized by Grb2 but these did not correspond to any known protein. Here we use recombinant Grb2 to isolate a complementary DNA called Gab1 (for Grb2-associated binder-1). Gab1 shares amino-acid homology and several structural features with IRS-1 (insulin-receptor substrate-1; refs 6,7), is a substrate of the EGF and insulin receptors, and can act as a docking protein for several SH2-containing proteins. Overexpression of Gab1 enhances cell growth and results in transformation. We conclude that Gab1 is a new protein in EGF and insulin receptor signalling which could integrate signals from different systems.

§ To whom correspondence should be addressed.

<sup>32</sup>P-labelled Grb2-fusion protein<sup>8</sup> was used to screen an expression cDNA library constructed from a glial tumour. We focused on one novel 1.6-kilobase (kb) clone. Rescreening of libraries resulted in the isolation of a 4.2-kb transcript which contained a polyadenylation signal and poly(A) tract. Sequence analysis revealed an initiator methionine followed by a 2.1-kb open reading frame. The conceptual translation predicts a 694-amino-acid protein of relative molecular mass 77K (Fig. 1a). A homology search showed Gab1 was most similar to human IRS-1, a major substrate of the insulin receptor<sup>6,7</sup>. The greatest homology (31% identity and 44% similarity) was in the pleckstrin-homology (PH) domain that is found in the amino terminus of both proteins<sup>9</sup> (Fig. 1b). The distal two-thirds of both proteins are proline- and serine-

rich, resulting in 47 predicted serine/threonine phosphorylation sites in Gab1, and 51 such sites in IRS-1 (Fig. 1c). An essential feature of IRS-1 is the presence of 20 potential tyrosine-phosphorylation sites that could recruit SH2 proteins enabling IRS-1 to act as a docking protein. It has been shown that Grb2 (ref. 10), phosphatidylinositol-3-OH kinase (PI(3)K; ref. 6), Nck (ref. 11), and SHPTP2/*syp* (ref. 12) directly interact with IRS-1. Gab1 contains 16 potential phosphotyrosine sites, suggesting that it could also function as a docking protein. Potential binding sites for the SH2 domains of PI(3)K, phospholipase C- $\gamma$  (PLC- $\gamma$ ), Nck and SHPTP2/*syp* were identified using consensus binding motifs<sup>13,14</sup>. IRS-1 contains a phosphotyrosine-binding domain<sup>15</sup> not found in Gab1, but the interaction with Grb2 provides an



**FIG. 1** Amino-acid sequence and distribution of Gab1. **a**, Amino-acid sequence of Gab1. Potential tyrosine-phosphorylation sites for known SH2 domains motifs are boxed<sup>13,14</sup>. These are: Grb2, Y-X-N-X, (Y48); Nck, Y-D-X-P, (Y242, Y307, Y406); phospholipase C- $\gamma$ , Y-X-I-P, (Y307, Y373, Y406); PI(3)K, Y-X-X-M (Y447, Y472, Y589); and SHPTP2/*syp*, Y-X-D-L (Y627). The SHPTP2/*syp* site is derived from the sequence in IRS-1 that binds this SH2 domain<sup>11</sup>. Two potential binding sites for the SH3 domains of Grb2 based on the motif P-X-X-P-X-R<sup>9</sup> are shown in bold. The GenBank accession number for Gab1 is U43885. **b**, Alignment between the pleckstrin-homology (PH) domains of Gab1 (amino acids 14–116) and human IRS-1 (amino acids 13–115). Hyphens denote gaps introduced to maximize the alignment, colons indicate amino-acid identity and stops denote similarity. The consensus sequence deduced for the six conserved subdomains of PH domains I–VI is shown below<sup>9</sup>. **c**, Potential phosphorylation sites in Gab1 and IRS-1 for casein kinase II (S/T-X-X-E/D), cAMP-dependent kinase (R/K-R/K-X-S/T), Cdc2 kinase (S/T-P-X-K/R), MAP kinase (P-X-S/T-P), and protein kinase C (S/T-X-R/K). Also shown are predicted binding sites for SH2 domains (Y), the PH domain, the proline/serine-rich region, binding sites for the SH3 domains of Grb2 (PRS), and the phosphotyrosine-binding region in IRS-1. **d**, Northern blots containing poly(A)<sup>+</sup> RNA from human tissues were hybridized with the Gab1 cDNA probe. Lanes: 1, spleen; 2, thymus; 3, prostate; 4, testes; 5, ovary; 6,

small intestine; 7, colon; 8, peripheral blood; 9, heart; 10, brain; 11, placenta; 12, lung; 13, liver; 14, skeletal muscle; 15, kidney; 16, pancreas. The migration positions of RNA size markers (in kb) is indicated on the left. **METHODS.** To clone Gab1, poly(A)<sup>+</sup> RNA from tumour D256 was primed using random hexamers. The cDNA was then cloned into the  $\lambda$ EX/ox expression vector (Novagen). The human Grb2 cDNA was cloned into a pGEX vector (Pharmacia) engineered to carry the phosphorylation site for protein kinase A (ref. 8). Fusion protein was labelled by *in vitro* phosphorylation and used at  $5 \times 10^5$  c.p.m. ml<sup>-1</sup> to screen the library. Oligo(dT) and hexamer-primed libraries were subsequently screened to obtain the complete cDNA. Sequencing reactions were performed using fluorescently labelled dideoxy terminators (Applied Biosystems) and run on an ABI373A automated sequencer. A multiple-tissue northern blot (Clontech) containing 2  $\mu$ g poly(A)<sup>+</sup> RNA per lane was hybridized and washed under standard conditions.

alternative means for Gab1 to associate with receptors. The similar number of tyrosine and serine/threonine phosphorylation sites in the much smaller Gab1 suggest that it is a compressed version of IRS-1.

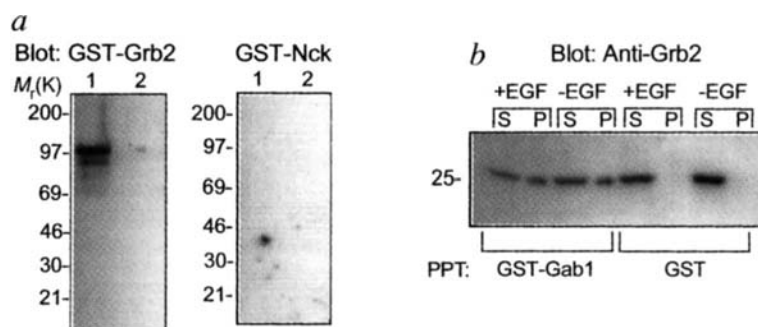
Northern blot analysis showed two Gab1 transcripts of 4.2 and 7.0 kb in all human tissues examined except for liver, lung and kidney (Fig. 1d). The complete open reading frame of Gab1 can be accounted for in the 4.2-kb transcript, so the 7.0-kb transcript may represent alternative splicing or a related gene. Northern blot analysis detects IRS-1 transcripts in skeletal muscle, heart and brain, although expression can be found in all human tissues by using the polymerase chain reaction with reverse transcription (RT-PCR)<sup>7</sup>. The easier detection of Gab1 transcripts suggests that it may be more prevalent than IRS-1.

Far-western blotting with Grb2 identified a ~100K protein in lysates from bacterial cells transformed with the proline/serine-

rich fragment of Gab1 which is not found in lysates from control cells (Fig. 2a). This recognition was specific because Nck, which contains three SH3 domains, did not bind the Gab1 fragment. To confirm that Gab1 interacts with native Grb2, a glutathione-S-transferase (GST) fusion protein was generated with this same Gab1 fragment (GST-Gab1) and used in precipitation experiments with A431 cell lysates. Grb2 was specifically bound by GST-Gab1 but not by GST alone (Fig. 2b). This binding did not require EGF, which was expected as SH3-domain interactions are growth-factor independent.

Anti-Gab1 immunoprecipitations were performed on A431 lysates and western blots were incubated with the same antibody. Endogenous Gab1 migrated as a ~115K protein band in unstimulated cells; this slower mobility may be due to a combination of its proline-rich nature and serine/threonine phosphorylation *in vivo*. EGF addition resulted in a further decrease in mobility to an

FIG. 2 Gab1 interacts with Grb2. **a**, <sup>32</sup>P-labelled fusion protein containing either Grb2 or Nck was incubated with western blots containing lysates from bacteria transformed with a plasmid encoding a fragment of Gab1 (lane 1) or no insert (lane 2). This fusion protein was ~20K greater than expected, which may be due to the high poline content. IRS-1 also migrates as a larger protein in SDS-PAGE<sup>9</sup> *M<sub>r</sub>* (in thousands) is indicated to the left. **b**, A431 cells were serum-starved and left untreated (-EGF) or stimulated with EGF (+EGF) and the cell lysates then used for precipitation with either GST-Gab1 or GST. Blots were incubated with an antibody against Grb2. S, supernatant; P, pellet from precipitations. The 25K size marker is indicated to the right. **METHODS**. ~2 µg of bacterial lysate from BL21 bacteria transformed with an *EXlox* plasmid containing amino acids 203-689 of Gab1, or no insert, was electrophoresed in SDS-PAGE. Western blots were incubated with 5 × 10<sup>5</sup> c.p.m. ml<sup>-1</sup> of either GST-Grb2 or a GST-fusion protein containing the entire coding region of human Nck. For GST-Gab1 precipitations, this same fragment was cloned into the pGEX vector. Quiescent A431 cells were left untreated or stimulated with EGF (100 ng ml<sup>-1</sup>) for 10 min. Cells were lysed as described<sup>30</sup>. 3 µg GST-Gab1 or GST was incubated with 150 µg A431



lysate for 30 min at 4 °C, and then incubated with 20 µl of glutathione beads (Pharmacia). Pellets were washed and one-third of the resulting pellet and supernatant was electrophoresed on SDS-PAGE. The western blots were incubated with an anti-Grb2 monoclonal antibody (Transduction Labs) and then with an <sup>125</sup>I-labelled anti-mouse secondary antibody.

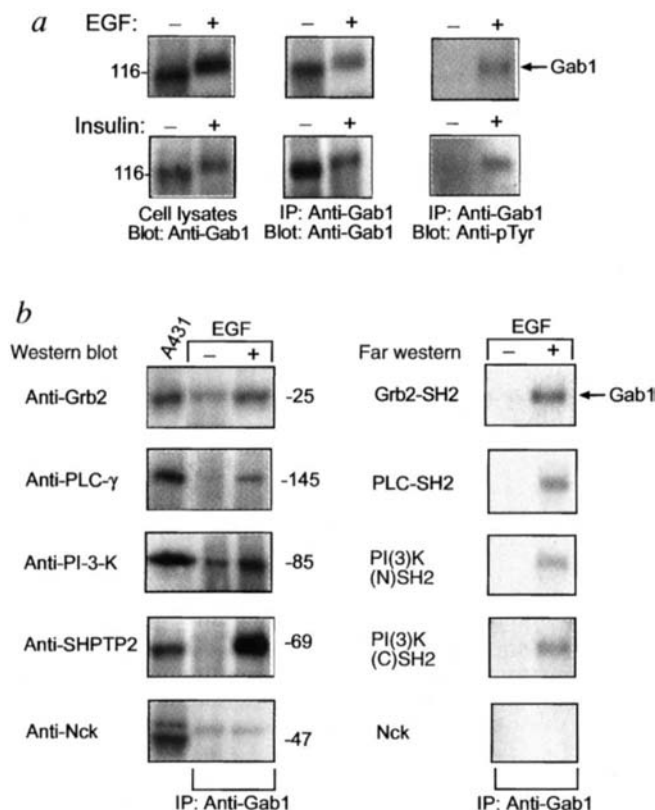


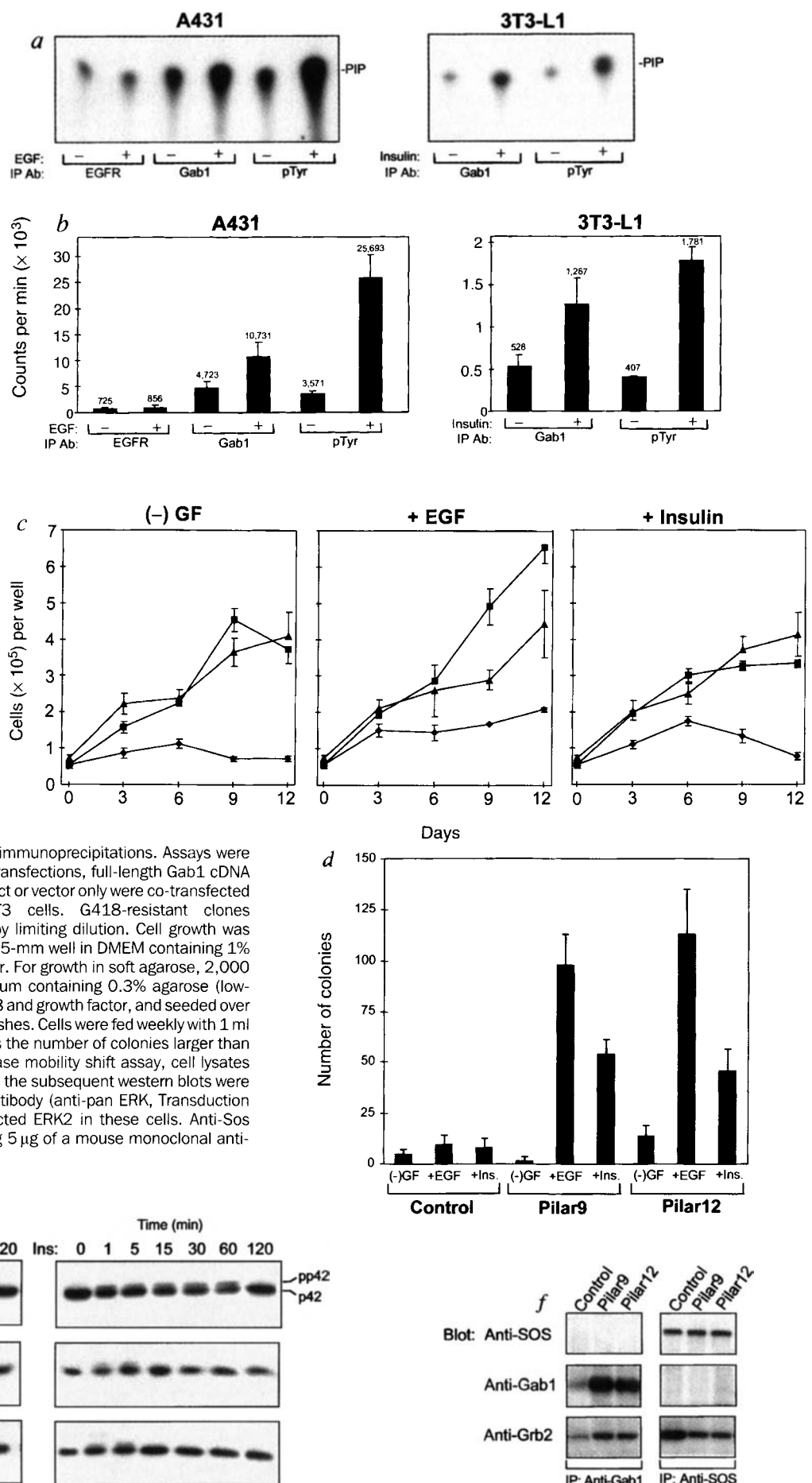
FIG. 3 Gab1 is a docking protein for the EGF and insulin receptors. **a**, Western blots containing lysates or anti-Gab1 immunoprecipitations from quiescent, EGF- or insulin-stimulated A431 cells were incubated with anti-Gab1 or anti-phosphotyrosine antibodies as indicated. *M<sub>r</sub>* (in thousands) is indicated to the left and the position of Gab1 is indicated on the right. **b**, Lysates from untreated or EGF-stimulated A431 cells were used for immunoprecipitation with anti-Gab1 antibody. Left panels: western blots were incubated with the indicated antibodies. Bands visible in the anti-Nck immunoprecipitation lanes represent crossreaction with the heavy chain of the anti-Gab1 antibody. We estimated that between 1-5% of the endogenous PLC-γ and PI(3)K, 5-10% of Grb2 and 10-15% of SHPTP2/syp associated with Gab1 following EGF addition. Right panels: portions of the blot containing Gab1 were incubated with GST-fusion proteins containing the SH2 domain from Grb2 (Grb2-SH2), both SH2 domains of PLC-γ (PLC-SH2), the N-terminal SH2 domain of PI(3)K (PI(3)K(N)SH2) or the C-terminal domain (PI(3)K(C)SH2), or total Nck.

**METHODS**. Quiescent A431 cells were left untreated, or stimulated with EGF for 10 min or insulin (100 nM) for 5 min. Cell lysate (50 µg) was run on the gel or 1 mg protein was used for immunoprecipitation with 5 µg of anti-Gab1 antibody. Western blots were incubated with anti-Gab1 antibody or anti-phosphotyrosine antibody (Upstate Biotechnology) followed by <sup>125</sup>I-labelled secondary antibody. Anti-Gab1 antibody was raised in rabbits by immunizing with GST-Gab1 and affinity-purified using GST-Gab1 coupled to agarose (AminoLink, Pierce). Antibodies against GST were removed by passage over a GST affinity column. Far-western blots were incubated with 2.5 µg ml<sup>-1</sup> GST-fusion protein, followed by anti-GST antibody (Santa Cruz Biotechnology), and then <sup>125</sup>I-anti-mouse antibody. The portions of the amino-acid sequence used were: Grb2-SH2, amino acids 50-161 of human Grb2; PI(3)K(N)SH2, amino acids 321-440, and PI(3)K(C)SH2, amino acids 614-724 of human PI(3)K; and PLC-SH2, amino acids 540-797 of human phospholipase C-γ<sub>1</sub>.



**FIG. 4** Gab1 mediates PI(3)K activity, cell growth and transformation. **a**, PI(3)K was assayed on immunoprecipitates (IP) produced by the indicated antibodies (Ab) on lysates from untreated, EGF-stimulated (A431) or insulin-stimulated (3T3-L1) cells. PIP, phosphatidylinositol-3-phosphate. **b**, Quantification of the PI(3)K assay. Bars represent average  $^{32}\text{P}$  c.p.m. values (Cerenkov counts) with s.d. Numbers above the bars designate average c.p.m. **c**, Cell-growth assays on the Pilar9 (squares) and the Pilar12 (triangles) cell lines (which express Gab1 at 13- and 8-fold over control (diamonds) cells, respectively) or a vector-only transfected cell line (control) in 1% serum without additional growth factor (-GF) or in the presence of EGF or insulin. **d**, These cell lines were seeded in soft agarose and grown in 10% serum without additional growth factor (-GF) or in the presence of EGF or insulin. **e**, Cell lines were serum-starved and then treated with EGF or insulin for the times indicated. Western blots were incubated with an antibody against MAP kinase (p42). Phosphorylated ERK2, which corresponds to the active form, is seen as a slower-migrating band (pp42). **f**, Lysates from the cell lines were immunoprecipitated with antibodies against Sos or Gab1. The subsequent western blots were incubated with antibodies against Sos, Gab1 or Grb2.

**METHODS.** PI(3)K was assayed as described<sup>18</sup> using 10  $\mu\text{g}$  anti-EGF receptor, 5  $\mu\text{g}$  anti-Gab1 and 15  $\mu\text{g}$  anti-phosphotyrosine antibody (PY20; Transduction Laboratories) for immunoprecipitations. Assays were run at least 5 times in duplicate. For transfections, full-length Gab1 cDNA was cloned into pLTR2 and this construct or vector only were co-transfected with pKOneo plasmid into NIH3T3 cells. G418-resistant clones (500  $\mu\text{g ml}^{-1}$ ) were subcloned twice by limiting dilution. Cell growth was assayed by seeding 50,000 cells per 35-mm well in DMEM containing 1% calf serum, G418 and the growth factor. For growth in soft agarose, 2,000 cells were suspended in 1 ml of medium containing 0.3% agarose (low-melting; Sigma), 10% calf serum, G418 and growth factor, and seeded over a 2-ml 0.6% agarose layer in 35-mm dishes. Cells were fed weekly with 1 ml suspension medium. After three weeks the number of colonies larger than 60  $\mu\text{m}$  was counted. For the MAP kinase mobility shift assay, cell lysates were electrophoresed in 8.5% gels and the subsequent western blots were incubated with an anti-MAP kinase antibody (anti-pan ERK, Transduction Laboratories) which consistently detected ERK2 in these cells. Anti-Sos immunoprecipitations were done using 5  $\mu\text{g}$  of a mouse monoclonal antibody (Transduction Labs).



apparent size of  $\sim 120$ K. An anti-phosphotyrosine antibody only detected the band in EGF-stimulated cells (Fig. 3a). Similar experiments showed that Gab1 was tyrosine-phosphorylated upon addition of insulin to A431 cells. *In vitro* kinase assays using EGF- or insulin-receptor preparations confirmed that GST-Gab1 was a direct substrate of these receptors (not shown).

Next we determined whether Gab1 was a docking protein for SH2 proteins. Western blots containing anti-Gab1 immunoprecipitations from unstimulated or EGF-stimulated A431 cells were incubated with antibodies against several SH2 proteins (Fig. 3b). Grb2, PI(3)K phospholipase C- $\gamma$ , and SHPTP2/syp were readily detected in these immunoprecipitations. Grb2 and PI(3)K showed an association with Gab1 in the absence of growth factor, but these associations were enhanced by EGF addition, suggesting further interaction via their SH2 domains. Grb2 can associate with Vav or c-Abl through its SH2 and SH3 domains<sup>3,4</sup>. The SH3 domains of Grb2 can associate with PI(3)K, which may explain the presence of this kinase in the immunoprecipitates from unstimulated cells<sup>16</sup>, or the SH3 domain of PI(3)K may bind to Gab1.

Far-western blots demonstrated that the SH2 domains from Grb2, PLC- $\gamma$  and PI(3)K, but not full-length Nck, bound to Gab1 from EGF-stimulated cells (Fig. 3b). Identical patterns of SH2-protein association were obtained from insulin-stimulated A431 cells (not shown). Thus, we have shown that tyrosine-phosphorylation of Gab1 mediates interaction with several proteins that contain SH2 domains. IRS-1 has similar associations except that it binds Nck but fails to bind PLC- $\gamma$  (ref. 17).

As IRS-1 is a major binding protein for PI(3)K in insulin signalling, we evaluated whether Gab1 serves a similar function in EGF signalling. As previously reported, EGF stimulation did not produce a substantial increase in PI(3)K activity in anti-EGF receptor immunoprecipitates (Fig. 4a, b)<sup>18</sup>. As expected, there was substantial PI(3)K activity in anti-Gab1 immunoprecipitates in unstimulated cells which increased 2.3-fold upon addition of EGF (Fig. 4a, b) and this was 42% of that found in anti-phosphotyrosine immunoprecipitates. In A431 cells, the EGF receptor can transphosphorylate ErbB3 and this constitutes  $\sim 49\%$  of the anti-phosphotyrosine-associated activity<sup>18</sup>, indicating that Gab1 and ErbB3 are two of the major PI(3)K-binding proteins after EGF stimulation. Insulin addition produced a 1.7-fold increase in Gab1 immunoprecipitates (results not shown). We also examined 3T3-L1 fibroblasts for insulin-stimulated activity. Although the magnitude of the response was lower, insulin produced a 2.4-fold increase in activity. This was 71% of that associated with anti-phosphotyrosine, indicating that Gab1 was the major binding partner in these cells following insulin addition. Interestingly, when 3T3-L1 are differentiated into adipocytes, PI(3)K activity is mainly associated with IRS-1 (ref. 19), suggesting a differentiation-state-specific role for these two docking proteins.

Both EGF and insulin induce mitogenesis and we asked if overexpression of Gab1 might augment cell growth. NIH3T3 cells were transfected with the Gab1 cDNA and two clones,

Pilar9 and Pilar12, were selected for further study. Both cell lines had an enhanced growth rate and achieved a high cell density in 1% serum, whereas control cells failed to proliferate (Fig. 4c). Growth factor addition did not significantly alter growth, except in the case of Pilar9, where EGF stimulated further growth at high density. These data suggest that Gab1 overexpression rendered NIH3T3 cells more sensitive to the limiting amount of growth factors present in serum. One measure of transformation is anchorage-independent growth. Both the Pilar9 and Pilar12 clones exhibited significant colony formation in soft agarose compared to control cells (Fig. 4d). This was dependent on growth-factor addition, indicating that activation of other proteins was necessary for transformation. Collectively, these data indicate that increased expression of Gab1 could facilitate tumorigenesis by receptor protein-tyrosine kinases.

As Sos and Grb2 can be coprecipitated with IRS-1, IRS-1 may participate in the activation of MAP kinase via the Ras pathway<sup>10,20</sup>. We tested whether the Gab1-transfected clones showed any enhancement of MAP kinase activation. Control cells had a large and sustained increase in MAP kinase phosphorylation following addition of growth factor (Fig. 4e), but the Pilar9 and Pilar12 clones showed only a small and attenuated duration of activation. We then investigated whether Sos could coprecipitate with Gab1. Whereas Grb2 was found in anti-Sos or anti-Gab1 immunoprecipitates, Sos could not be detected in anti-Gab1 complexes, nor could Gab1 be found in anti-Sos complexes (Fig. 4f). The amount of Grb2 that coprecipitated with Sos was decreased in the Pilar9 and Pilar12 clones, so an explanation for the decrease in MAP kinase activity is that Gab1 competes with Sos for binding to Grb2. Although it is not clear whether Gab1 regulates Sos and the Ras pathway under normal conditions, these results show that overexpression of Gab1 does not enhance MAP-kinase activation and that Gab1 and Sos form separate complexes with Grb2. These findings are consistent with data showing that IRS-1 is not directly involved in MAP kinase activation by insulin<sup>21-23</sup>.

The identification of Gab1 shows that a docking protein is also present in the EGF-receptor pathway and suggests that other such proteins may be found for other receptor protein-tyrosine kinases. Through Grb2, Gab1 could also be involved in the pathways of other tyrosine kinases. For example, Grb2 binds to focal-adhesion kinase (FAK) and mediates signalling by integrins<sup>24</sup>. Insulin promotes the association of IRS-1 with  $\alpha_v\beta_3$ , but this interaction appears to be specific for this integrin<sup>25</sup> so a FAK-Grb2-Gab1 complex may be an alternative means to initiate signalling. IRS-1-knockout mice show growth retardation but few other abnormalities<sup>26,27</sup> because of the compensatory expression of another docking protein, IRS-2 (ref. 28), suggesting that docking proteins are essential for insulin signalling. It will be interesting to see whether Gab1 is upregulated in these mice and to establish the contribution of Gab1 to various aspects of cell signalling and disease processes. □

Received 7 September; accepted 6 December 1995.

- Lowenstein, E. J. *et al. Cell* **70**, 431–442 (1992).
- Downward, J. *FEBS Lett.* **338**, 113–117 (1994).
- Ren, R., Ye, Z. S. & Baltimore, D. *Genes Dev.* **8**, 783–795 (1994).
- Ye, Z. S. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12629–12633 (1994).
- Gout, I. *et al. Cell* **75**, 25–36 (1993).
- Sun, X. J. *et al. Nature* **352**, 73–77 (1991).
- Araki, E. *et al. Diabetes* **42**, 1041–1054 (1993).
- Ron, D. & Dressler, H. *Biotechniques* **13**, 866–869 (1992).
- Musacchio, A., Gibson, T., Rice, P., Thompson, J. & Saraste, M. *Trends biochem. Sci.* **18**, 343–348 (1993).
- Skolnik, E. Y. *et al. Science* **260**, 1953–1955 (1993).
- Lee, C. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 11713–11717 (1993).
- Sun, X. J., Crimmins, D. L., Myers, M. G., Miralpeix, M. & White, M. F. *Molec. cell. Biol.* **13**, 7418–7428 (1993).
- Songyang, Z. *et al. Molec. cell. Biol.* **14**, 2777–2785 (1994).
- Songyang, Z. *et al. Cell* **72**, 767–778 (1993).
- Gustafson, T. A., He, W., Craparo, A., Schaub, C. D. & O'Neill, T. J. *Molec. cell. Biol.* **15**, 2500–2508 (1995).
- Wang, J., Auger, K. R., Jarvis, L., Shi, Y. & Roberts, T. M. *J. biol. Chem.* **270**, 12774–12780 (1995).
- Myers, M. G. & White, M. F. *Diabetes* **42**, 643–650 (1993).

- Soitoff, S. P., Carraway, K. L., Prigent, S. A., Gullick, W. G. & Cantley, L. C. *Molec. cell. Biol.* **14**, 3550–3558 (1994).
- Lamphere, L., Carpenter, C. L., Sheng, Z.-F., Kallen, R. G. & Lienhard, G. E. *Am. J. Physiol.* **266**, E486–494 (1994).
- Baltensperger, K. *et al. Science* **260**, 1950–1952 (1993).
- Yamauchi, K. & Pessin, J. E. *J. biol. Chem.* **269**, 31107–31114 (1994).
- Pruett, W. *et al. Molec. cell. Biol.* **15**, 1778–1785 (1995).
- Chen, D., Van Horn, D. J., White, M. F. & Backer, J. M. *Molec. cell. Biol.* **15**, 4711–4717 (1995).
- Schlaepfer, D. D., Hanks, S. K., Hunter, T. & van der Geer, P. *Nature* **372**, 786–791 (1994).
- Vuori, K. & Ruoslahti, E. *Science* **266**, 1576–1578 (1994).
- Tamamoto, H. *et al. Nature* **372**, 182–186 (1994).
- Araki, E. *et al. Nature* **372**, 186–190 (1994).
- Sun, X. J. *et al. Nature* **377**, 173–177 (1995).
- Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. *Science* **266**, 1241–1247 (1994).
- Moscattello, D. K. *et al. Cancer Res.* **55**, 5536–5539 (1995).

ACKNOWLEDGEMENTS. We thank K. Kinzler and L. K. Su for advice on library screening, D. Lawe for reading the manuscript, and H. Alder of the Nucleic Acid Facility of the Jefferson Cancer Institute for expert work. This research was supported by grants from the NIH and the Ronald McDonald Children's Charities.

# Biotechnology bulletin

**For scaling-up bioprocess operations and ensuring quality control there are a number of new products described below including a chromatography control system, an on-line biochemical analyser and a non-invasive animal testing service.**

Antisoma has obtained an exclusive world-wide licence from the UK National Institute of Cancer Research for a **monoclonal antibody against a marker expressed uniquely on the blood vessels of tumours**. When a tumour grows in the body, it pirates nearby blood vessels causing a blood supply to be plumbed into the tumour, thus providing the cancer cells not only with nutrients but also a route to spread throughout the body. This antibody has the ability to target the marker of tumour angiogenesis and potentially cut off the tumour's blood supply. The manufacturer states that it may prove useful for not only detecting cancer, as shown in early pilot clinical studies, but may also destroy it. It is also being tested in other abnormal situations, for example in inflammation where new blood vessels may develop, or in cases of blindness where abnormal blood vessels grow on the retina.

**Reader Enquiry No. 100**

InVivoMetrics now offers non-invasive **techniques to establish the safety and efficacy of new drugs or medical devices**. The services should prove useful to manufacturers in the preclinical phase of development and should assist in satisfying the US Food and Drug Administration's requirements to demonstrate safety and efficacy before human clinical testing. This technology is designed to collect information without harming or interfering with the physiology of test animals. It is said to enable repeated observations of the same animal, thus using far fewer animals (up to 90 per cent fewer). Magnetic resonance imaging (MRI) is used to provide detailed information about the bones, eyes, blood vessels and internal organs of an intact organism. MRI can visualize structures as small as 0.1 mm so that changes can be tracked in the internal structure and function of organs, and evaluate their response to drugs. Nuclear magnetic resonance spectroscopy (NMR) is used to evaluate the chemical properties inside living animals, such as intracellular pH, metabolism and blood oxygen levels without dissecting the animal.

**Reader Enquiry No. 101**

## Quality assurance

The new Hewlett-Packard HP 1100 series **HPLC systems and modules combine validation and certification-support**

capabilities with design, manufacturing and support features. The product family consists of an isocratic pump, a binary pump, a quaternary pump, a vacuum degasser, a manual injector, an autosampler, a thermostatted column compartment, a variable wavelength detector, a diode-array detector, a control module and an HP ChemStation for HPLC. A typical QA/QC laboratory manager or trained technician can develop a method and store it on an insertable PC card. A less-skilled operator can then plug the

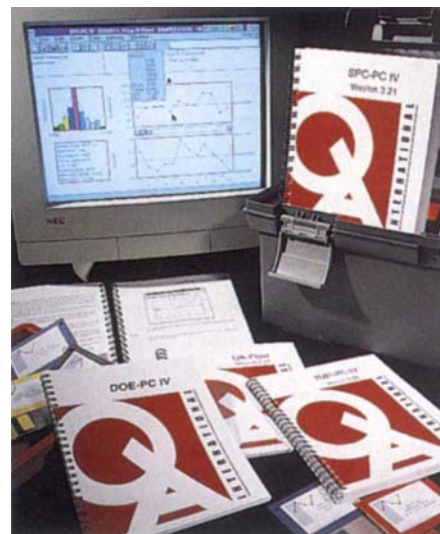


**Hewlett-Packard 1100 series HPLC.**

PC card into the instrument to set up the method automatically. Validated methods and procedures can be sent to geographically distant laboratories by e-mail, thereby providing consistent, error-free method loading every time.

**Reader Enquiry No. 102**

A suite of **software tools for quality-management, analysis and control** has been introduced by Adept Scientific. The QASI range comprises seven new quality-management and planning tools, all seven traditional quality-control tools, as well as programs for the design of experiments, calibration, flowcharting and interactive SPC training. The new tools are available in the 7MP-PC IV package, which is designed to take the guesswork out of planning and implementing a quality project. These tools can help recognize and organize information trends and patterns; highlight relationships between items of information and identify underlying causes and effects; systematically



**Adept scientific QA software with new tools for quality analysis and control.**

weigh the comparative significance of specific assignments, tasks and options; rate the correlation between groups of ideas and identify the critical ones; analyse the steps of a proposed process and look at the potential changes; break down broad goals or ideas into specific actions; structure the process; determine the critical path to follow; and establish the time needed to accomplish all tasks and sub-paths.

**Reader Enquiry No. 103**

## Separation techniques

PerSeptive Biosystems has introduced the Scout column selector, an **automated multi-column switching device** for use with the company's Windows-based BioCAD and BioCAD/Sprint perfusion chromatography systems in the purification of biological molecules. The column selector extends the purification capability of these systems by automating and accelerating the screening of multiple surface chemistries to determine the optimal separation conditions for specific biomolecules. There are 25 different surface chemistries for the Poros perfusion chromatography media. The device is available as an option on all new BioCAD systems, or as an upgrade.

**Reader Enquiry No. 104**

Pharmacia Biotech has introduced Unicorn version 2.0, a control system with the ability to network all the company's chro-



matography systems from bench-scale up to full-scale production. The system is intended for method development, clear presentation of system status and automatic data evaluation and it also has report generation capabilities. Using a single, uniform software system for all scales of purification should save on costs, minimize operator training time and improve efficiency. Validation support software allows users to monitor and control all systems from any PC connected to the network and access data for written reports. The software also includes security features to prevent unauthorized access to systems and data.

**Reader Enquiry No. 105**

Molecular Dynamics' new **Storm system for gel and blot analysis** unites storage phosphor autoradiography, fluorescence and chemifluorescence detection in a single instrument. It is designed for researchers who wish to use emerging fluorescence-based gel and blot imaging techniques, while continuing to use autoradiography. The variable-mode imaging architecture integrates the optical components required for the system's three-in-one capability into a single-scanning element. The appropriate components are automatically activated when the user selects one of the three scanning modes from the system's computer screen. Storm has a 35 cm × 43 cm scan area to accommodate sequencing gels and other large samples. ImageQuant software for Macintosh or PC computers provides system control and data analysis capabilities.

**Reader Enquiry No. 106**

## Biochemical analysis

The new YSI 2700 Select **biochemistry analyser** can monitor and control bioprocesses aseptically with no risk of contamination. With the addition of the YSI 2730 monitoring and control accessory, the analyser becomes an on-line monitoring system that maintains the sterility of the process. The analyser pulls samples directly from the process vessel, as often as every two minutes around the clock, while maintaining sterility in the system. The analyser will then automatically replenish nutrients or add diluent when concentrations vary from the setpoints. As a single-channel unit, the analyser provides 60-s measurements of glucose, lactate, glutamate, ethanol, sucrose, lactose, galactose, choline or hydrogen peroxide. It uses a thin film of enzyme immobilized within a membrane to produce hydrogen peroxide from substrates of the enzyme. The hydrogen peroxide is measured electrochemically at a platinum anode, and as the enzymes are immobilized they can be used many times.

**Reader Enquiry No. 107**

The ConfoCor system examines molecular interactions and should find application in the fields of pharmaceutical research, molecular biology and immunology. Jointly developed by Carl Zeiss and Evotec BioSystems, the system uses the principle of fluorescence-correlation spectroscopy (FCS). It employs Zeiss microscope optics and signal processing equipment. The Axxess-to-FCS software package from Evotec is designed to enable exact, quasi real-time determination of molecule parameters, such as diffusion time, binding constant, concentration and reaction-kinetic constant.

**Reader Enquiry No. 108**

## Publications

The 1995 edition of the *UK Biotechnology Handbook* has now been published by BioCommerce Data. Now in its sixth edition, it lists over 720 British organizations involved in biotechnology and contains over 2,900 senior contact names. Each listing has been updated and verified for the new edition, which has over 135 new entries. The handbook includes review articles by industry experts on topical subjects, such as patent laws in Europe and the US, financing, recruitment, the Department of Trade and Industry's Biotechnology Means Business initiative, risk assessment, bioinformatics and the Internet. Also, the company has just launched a new Windows version of its international biotechnology business information service on CD-ROM, Bio Commerce Abstracts. It is designed for users with no previous experience using information retrieval tools and provides concise summaries of over 300,000 news articles. It can be used on either IBM PCs (running SoftWindows 2.0) or Macintosh machines.

**Reader Enquiry No. 109**

Knight-Ridder Information has released a new **CD-ROM product: *Biotechnology and BioEngineering***. This database provides comprehensive, current information and is compiled from over 2,000 primary scientific publications, including journal articles, monographs, conference proceedings, technical reports and books, with coverage from 1984 to the present. The database covers new drug developments and delivery systems, new techniques for water treatment and industrial hygiene, advances in plant breeding and biopesticides, patents for new products, the latest developments in important scientific projects, advances in medical instrumentation, imaging, biomechanics and researching the latest applications in energy production and conservation.

**Reader Enquiry No. 110**

*These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.*

## ADVERTISEMENTS

### LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS — EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

**DIANDROM**

P.O. Box 650126, D-81215 Munich  
For USA call 1-800 Liposome

**READER ENQUIRY NO. 14**

### Menai Organics Limited

Unit 5, Menai Technology Centre, Deiniol Road, Bangor, Gwynedd, N. Wales LL57 2UP.

Telephone (01248) 354103 Ext 205

### RARE ORGANIC CHEMICALS FOR LABORATORY USE AND CUSTOM SYNTHESIS

Contact us now for your free copy of our catalogue of screening compounds, also available on 3.5" disk.

**READER ENQUIRY NO. 15**

NEWCASTLE UNIVERSITY  
MOLECULAR BIOLOGY UNIT



### PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

**£12 per cycle**

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray  
Phone: +44 191 222 8612  
Fax: +44 191 222 8129  
Email: JOE.GRAY@newcastle.ac.uk

**READER ENQUIRY NO. 12**

### OLIGONUCLEOTIDES

**£1 per base**

- stringent quality control
- express turnaround
- column purified: ready to use

*By research labs for research labs*

Molecular Biology Services

King's College School of Medicine and Dentistry  
Rayne Institute, 123 Coldharbour Lane, London SE5 9NU  
Tel: 0171-346 3126 Fax: 0171-733 3877

**READER ENQUIRY NO. 10**